2014

Visualizing and Predicting the Effects of Rheumatoid Arthritis on Hands

Radu P. Mihail
University of Kentucky, r.p.mihail@gmail.com

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Radu P. Mihail, Student

Dr. Nathan Jacobs, Major Professor

Dr. Miroslaw Truszczynski, Director of Graduate Studies

Visualizing and Predicting the Effects of Rheumatoid Arthritis on Hands

---

DISSERTATION

A dissertation submitted in partial
fulfillment of the requirements for the
degree of Doctor of Philosophy in the
College of Engineering at the
University of Kentucky

By
Radu P. Mihail
Lexington, Kentucky

Co-Directors: Dr. Nathan Jacobs and Dr. Judy Goldsmith
Professors of Computer Science
Lexington, Kentucky 2014

ABSTRACT OF DISSERTATION

Visualizing and Predicting the Effects of Rheumatoid Arthritis on Hands

This dissertation was inspired by difficult decisions patients of chronic diseases have to make about about treatment options in light of uncertainty. We look at rheumatoid arthritis (RA), a chronic, autoimmune disease that primarily affects the synovial joints of the hands and causes pain and deformities. In this work, we focus on several parts of a computer-based decision tool that patients can interact with using gestures, ask questions about the disease, and visualize possible futures. We propose a hand gesture based interaction method that is easily setup in a doctor's office and can be trained using a custom set of gestures that are least painful. Our system is versatile and can be used for operations like simple selections to navigating a 3D world. We propose a point distribution model (PDM) that is capable of modeling hand deformities that occur due to RA and a generalized fitting method for use on radiographs of hands. Using our shape model, we show novel visualization of disease progression. Using expertly staged radiographs, we propose a novel distance metric learning and embedding technique that can be used to automatically stage an unlabeled radiograph. Given a large set of expertly labeled radiographs, our data-driven approach can be used to extract different modes of deformation specific to a disease.

KEYWORDS: RA, deformities, visualization, hands, xrays

Radu P. Mihail
Date: 05/02/2014

Visualizing and Predicting the Effects of Rheumatoid Arthritis on Hands

By
Radu P. Mihail

Co-Directors of Dissertation: Dr. Nathan Jacobs and Dr. Judy Goldsmith

Director of Graduate Studies: Dr. Miroslaw Truszczynski

Date: 05/02/2014

For my wife, Marina.

ACKNOWLEDGMENTS

This dissertation would not be possible without the help of many people throughout the years. I was fortunate to have been advised by two amazing people, who put up with me for the duration of my graduate studies and helped minimize failure modes. Thank you Judy Goldsmith and Nathan Jacobs for the cool courses I took from both of you, for the encouragement I received, for the advice you've given, for the immaculate professionalism you showed during the years, and for the support to pursue my interests. This dissertation would not have been possible without you.

I take this opportunity to thank the rest of my committee, Ruigang Yang, Sen-ching (Samson) Cheung, Ruigang Yang and Kristine Lohr for their invaluable comments and feedback over the years. I would like to thank Kristine Lohr and Gustav Blomquist for everything, especially the help with the medical IRB and the invaluable radiographs used throughout this work.

My immediate family has been there for me during the ups and downs, thanks for the help you've given along the way to keep me in touch with my better self, Lili, Walter—may your gentle soul rest in peace, Bebi and Adi. Thank you Unchiu' (Adrian) for help on countless occasions. Thank you Randy and Lauren for your support and believing in me.

Last and by far most to Marina, for everything.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1 Introduction

1.1 Motivation

This dissertation is motivated by the difficult decision making processes that patients facing a recent diagnosis of a chronic disease must make. In particular, I focus on rheumatoid arthritis (RA), a chronic inflammatory autoimmune disease that affects the musculoskeletal system, primarily causing deformities of the hands and pain. The work presented in this dissertation was inspired by game-based decision aids which take advantage of an interactive environment, distilled medical data and clever risk communication strategies to inform patients of their disease and available treatment choices. In a standard shared medical decision setting, where patients are expected to take an active role in their health care, several complicating factors arise:

- Health care economics limit patient-practitioner interaction time.
- Disorder is chronic and the symptoms are difficult to explain.
- Treatment options are complex, with uncertain benefits/side-effects.

This dissertation is the result of attacking a subset of problems associated with engineering a decision tool that can allow patients to safely explore the effects of RA on their hands.

1.2 Main Contributions and Related Publications

The contributions shown in this work were inspired by an interactive game-based patient decision aid (DA) intended to help RA patients understand their disease and treatment options. Most of the work was done jointly with other researchers, thus I use "we" when describing this work. The coauthors played a critical role in the development of this dissertation and I owe much of my success to my advisors: Drs. Judy Goldsmith and Nathan

Jacobs, medical doctors: Drs. Kristine Lohr and Gustav Blomquist, and two undergraduates: Kaitlin Burton and Frederick Hallock, whom I supervised to write C# code and a resulting paper about visualization of RA damage to hands.

This dissertation presents contributions in the fields of human-computer interaction (HCI), computational vision, medical information visualization and machine learning. I focused on three aspects involved in building a game-based decision tool for RA patients:

- Human-Computer Interaction: RA patients have difficulties using a standard keyboard and mouse due to deformities and pain. To this end, we propose a novel touchless hand gesture recognition system that can be used by patients to perform actions as simple as selecting an item from a menu to navigating an avatar through a 3D world.

- Computational Vision: Radiographs (X-Rays) are commonly used as supporting evidence in the diagnosis of RA. We propose a novel point distribution model (PDM) that is capable of capturing two important destructive effects of RA on the synovial joints in the hand: joint space narrowing and subluxation (shifting and dislocation). We also present a novel formulation of fitting the PDM to radiographs as inference in a conditional random field (CRF).

- Information Visualization: We show how our proposed PDM can be used to animate a 3D hand model using forward kinematics (FK). Using techniques from computer graphics, we show novel visualizations of RA disease progression.

- Machine Learning: We propose a novel dimensionality reduction technique that exploits disease stage as a discrete label. This work is inspired by the observation that variance due to the disease in healthy samples should be low, while variance due to progressive changes due to disease is high. To this end, we formulate new dimensionality reduction techniques that cluster healthy samples near the origin and more advanced stage samples away from the origin. We show how these models can be

used to sample progression paths conditioned on a series of radiographs from the same patient.

Human-Computer Interaction (HCI)

In the work on HCI we proposed a novel static hand gesture recognition system that utilizes two Kinect sensors. The sensors are placed on the front left and right sides of a sensing area, and can easily be set up in a doctor's office. We extract a rich point cloud from the Kinects, out of which gestures and hand orientation are detected in real-time. We compute histogram-based descriptors from the point cloud and employ the majority rule voting scheme to detect the gesture. We achieve rotation invariance by using the forearm as a cue for hand orientation and align the hand with the world coordinate system. We evaluate our system under different and challenging motion and rotation conditions. This work has resulted in a joint conference publication:

- Radu P. Mihail, Nathan Jacobs and Judy Goldsmith, "Real Time Gesture Recognition With 2 Kinect Sensors", 16th International Conference on Image Processing, Computer Vision, & Pattern Recognition (IPCV), 2012.

Computational Vision

Postero-anterior (PA) view radiographs are a standard imaging technique used in the diagnosis and evaluation of RA disease progression. Since standard radiographs are relatively inexpensive, many rheumatologists choose to order them. Dr. Gustav Blomquist of the Radiology Department at the University of Kentucky Chandler Hospital was generous to provide us with a unique dataset of radiographs. He used his expertise to label them in 4 stages: healthy, early, moderate and late. We show a novel way to extract information pertinent to the effects of the disease on hands visible in radiographs. Our proposal is a novel point distribution model (PDM), that is capable of capturing joint space narrowing and subluxation, which are key indicators of progressive disease activity. We propose a

novel formulation of fitting the PDM to radiographs as inference in a conditional random field (CRF). We engineered potential functions that focus on specific anatomic structures and a learned shape prior. We evaluate our approach on the UK Radiology dataset as well and a publicly available dataset, as well as show the relative importance of each potential function. We show the usefulness of this approach as an initialization step to automatically estimate bone contours. This work has resulted in a joint conference publication:

- Radu P. Mihail, Gustav Blomquist and Nathan Jacobs, "A CRF Approach to Fitting a Generalized Hand Skeleton Model", Winter Conference on Applications of Computer Vision (WACV), 2014.

Information Visualization

When patients plan on the treatment choice with their health care provider, they often ask about the trade-offs, as well as the consequences of selecting no treatment at all. The option to choose no treatment should come to the reader as no surprise, since many of the drugs come with scary, albeit low probability side-effects. This dilemma inspired us to create a visualization of the progressive damage due to RA on hands. We use our proposed PDM to deform a hand mesh model using standard graphics techniques. To the best of our knowledge, this is the first attempt to visualize damage caused by RA using computer graphics in a data-driven manner. Parts of this work were done under my supervision by two undergraduate students, Kaitlin Burton and Frederick Hallock. This resulted in a joint conference publication:

- Kaitlin Burton, Frederick Halloc and Radu P. Mihail, "A Data-Driven Approach to Visualize the Effects of Rheumatoid Arthritis on Hands", International Conference on Computer Games (CGAMES), 2013.

Machine Learning

Dimensionality reduction techniques are useful as a visualization tool for high dimensional

data (e.g., when the target dimensionality is 2 or 3) and also to improve algorithms which rely on distance computations such as nearest neighbor, support vector machines and unsupervised clustering algorithms. Similarity measures (i.e., distance metrics) provide insights into the underlying structure of the data and help build better visualizations through embeddings. In the medical image analysis field, data are often extracted as shapes, an example of which are PDMs. Of particular interest to us are diseases that show progressive damage, such as RA. Using our PDM as input data, we propose an algorithm to learn a quadratic Gaussian metric that exploits the progressive nature of RA. Our algorithm is based on the intuition that variability exists due to two separate factors: differences between patients and differences due to the progressive damage from RA. The geometric intuition is that PDMs of healthy hands should **ideally** collapse to the origin of the embedding space, while the PDMs with signs of RA should be mapped at a distance from the origin *proportional* to the stage at which the PDM was captured. We formulate the optimization framework with the stage constraints and demonstrate two use cases: classification (e.g., stage given a PDM) and as progression models given a series of 2 or more PDMs of the same patient.

1.3 Structure of the Dissertation

In this dissertation I will walk the reader through a set of computational tools we developed and their respective contributions to the field of computer science, inspired by decision making processes for victims of chronic diseases facing complex treatment decisions.

Chapter 2, Background: This Chapter is intended for readers not familiar with computer science or mathematics. I will provide the background information necessary to understand this work as it fits in the areas of computer science where this dissertation makes contributions: human-computer interaction (HCI), computer vision, visualization and machine learning. In Section 2.1 I describe rheumatoid arthritis and some essential anatomical terminology useful to understand this work. More specifically the medical image analysis contribution in Chapter 5 is motivated by key anatomical

changes specific to RA. In Section 2.2 I provide a brief survey of HCI and existing methods that focus on individuals with disabilities that prevents normal use of a keyboard and mouse as standard input. In Section 2.4 I give an overview of the field of machine learning and detail specifics relevant to Chapters 4, 5 and 6. In Section 2.3 I give a brief overview of conditional random fields.

Chapter 3, Decision Aids for RA Patients: This chapter is a review of existing patient decision aids for patients with RA. I focus on challenges involved in the design of good DAs such as risk and numerical information delivery, available treatments for RA, difficulties patients face when choosing a treatment, available DAs for DMARDs and biologics and I finally conclude with suggestions on DA design and how gaming technology can improve the decision tools.

Chapter 4, Kinect Based Hand Gesture Recognition: In this chapter I describe the work we have done on HCI and how it applies to our target audience of RA patients. I describe the gesture recognition system in Section 4.1 and a comprehensive evaluation with challenging hand movement and rotation in Section 4.3.

Chapter 5, A CRF-based Approach to Fitting Generalized Hand Skeleton Models: In this chapter I detail the work on fitting a model to radiographs of hands in the postero-anterior view. We describe a new point distribution model capable of capture key anatomical deformations caused by RA and formulate the fitting problem as inference in a conditional random field (CRF). We evaluate our model on two datasets and show how it can be used to initialize a contour fitting algorithm to segment bones, a challenging problem of interest to RA disease progression assessment.

Chapter 7, Visualizing Hands Affected by RA: This chapter describes a visualization technique for RA progression using the PDMs described in Chapter 5. We show how to rig a hand model and animate a progression from healthy to late stage RA.

Chapter 6, Disease Stage Metric Learning: In this chapter, we propose a novel metric learning algorithm based on the medical intuition that anatomical shape variability comes from two sources: differences between patients and monotonically increasing shape variability due to destructive changes from a disease. The metric learned is a quadratic Gaussian (Mahalanobis) distance that we optimize to embed healthy samples close to the origin of the latent space, while diseased hands proportionally far from the origin.

Chapter 8, Conclusions and Future Directions: In this chapter we look back at the contributions of this dissertation in the context of building blocks for a game-based DA. I will outline directions for future research that will lead to better decision making tools for patients with chronic conditions.

Chapter 2 Background

In this chapter, I will attempt to give my readers the appropriate background information necessary to understand the entire dissertation. I begin with a an overview of rheumatoid arthritis and anatomical terminology. This disease motivated and inspired a large part of this work and is discussed in detail in Section 2.1. I then provide a brief overview of HCI for the disabled in Section 2.2, preliminary information needed to understand the technical contributions in mathematics for machine learning in Sections 2.4 and Section **??**.

Our long term goal is to create a data-driven patient decision aid based on solid strategies for risk communication, visualization and medical knowledge about the progressive destructive effects on the musculoskeletal system. This disease affects more than the joints in the hands, but our hands are very useful appendages to perform daily activities. When this functionality is lost over time due to RA, it leads to disabilities and a societal cost that can, to some degree, be prevented.

When I worked with colleagues from the Kentucky Clinic and University of Kentucky's Chandler Hospital to collect radiographs, I asked why the sample of late stage RA radiographs was smaller. The answer was simple: there are better treatments and patients do not get to that point as they did even a decade ago. This is encouraging, but depends on patients taking an active role in their health care.

The practice of medicine evolved from a paternalistic model, where the physician is seen as the expert, dominating the medical encounter and using his/her skills to recommend tests/treatments, to a shared decision-making model where the key idea is that informed patients are better patients [19]. Patients facing a serious illness and uncertainty regarding the outcome are vulnerable to psychological and physiological stress. In these instances, patients should be given a systematic way to structure the decision-making process, i.e., as Charles et al. [19] put it:

> "... efforts to promote shared decision-making may well require interventions that not only provide patients with information but also with a way of thinking about treatment decision-making that helps them focus on key issues and evaluate relevant options. "

Due to a harsh health care economic reality in which patient-physician interaction times are limited to under 10 minutes, patients are left unempowered, lacking key information and a systematic way to evaluate their options and relevant uncertain outcomes. It is here where computer-based patient decision tools have the potential to make a positive impact on patients by empowering them with information and strategies.

This dissertation is the result of addressing problems RA patients encounter shortly after the diagnosis. They face an unforgiving disease that left untreated can leave them disabled, deformed, depressed and in pain. Often, patients visit a rheumatologist long after the onset of the disease, time at which some deformity already occurred and pain can be disabling. Pain and existing deformity can prevent them from using a keyboard and mouse to perform simple tasks on a computer, e.g., read informational materials on the web. To address this challenge, we developed a touchless hand gesture recognition system that is highly customizable and easily setup in a doctor's office, described in detail in Chapter 4.

When patients plan, they have questions about the disease and its effects on their well-being. Deformities due to RA are a function of many variables and vary with respect to treatment and time. In this context, I asked the following question: can I show patients how their hands *might* look like in the future? A trivial solution would be to show photos of other patient's hands. In search of a better solution, I quickly realized that visualization of possible progression paths could be a powerful tool for patients making treatment decisions.

With the help of rheumatologist Dr. Kristine Lohr and radiologist Dr. Gustav Blomquist, I collected a set of radiographs of patients hands in postero-anterior view, where each radiograph was skillfully assigned one of four labels: healthy, early, moderate and late. We

needed to extract relevant information from these radiographs, i.e., a compact representation of joint damage due to RA. To this end, we developed a novel point distribution model (PDM) that captures two key features of RA deformities: joint space narrowing and subluxation (dislocation). We describe this problem in Chapter 5, in the context of machine vision and propose a novel formulation of the problem to automatically fit the PDM.

We take advantage of this compact representation to solve two problems:

- Learn a model of disease progression, described in detail in Chapter 6.
- Visualize damage visible in radiographs due to RA, in Chapter 7.

In the remainder of this chapter, I will give the reader a thorough description of the background needed to understand this thesis.

2.1 Rheumatoid Arthritis

RA affects many tissues in the body, but it has particularly evident effects on the synovial joints (see Figure 2.1). The synovium, also called the *stratum synoviale*, is a soft tissue that surrounds the articular tissue in synovial joints. Symptoms of RA can vary in intensity and frequency, but with high probability, destruction and deformity of joints occurs in parallel with lethargy and depression. The progression of the disease is unpredictable as it can potentially simmer for years with undetectable symptoms, then flare up aggressively followed by a potential abrupt flare down. On average, the disease progresses punishingly and leaves its victims with deformed extremities, depressed and disabled.

RA does not have a one source cause; instead, experts [67, 65] believe the inflammatory responses of the body manifests as a result of a set of factors and a likely genetic susceptibility component. Weakened immune systems, internal and external allergic response to unknown allergens, pathogens (e.g., candidiasis), inappropriate nutrition, hormonal imbalances, stress, exposure to pollution as well as other unknown factors.

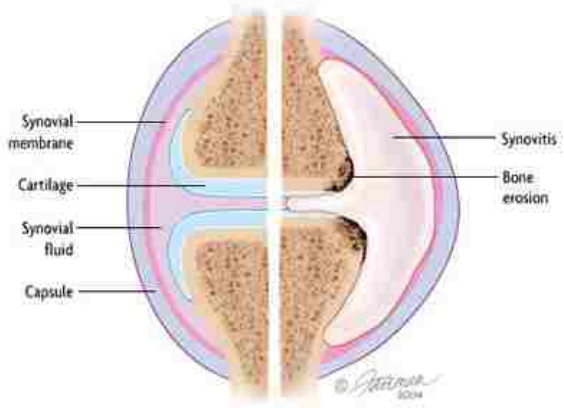


Figure 2.1: Image retrieved from the American College of Rheumatology. On the left one can see the normal structure of a joint. Due to autoimmune reactions caused by RA, tissues such as the cartilage and bone erode. Synovitis also causes swelling and pain.

At the time of this writing, RA does not have a cure. The American College of Rheumatology (ACR) published recommendations for the use of disease-modifying antirheumatic drugs and biologic agents in the treatment of RA [93]. These guidelines are intended to guide therapy. In light of uncertainty and factors such as drug interactions, conflicting information from clinical trials and patient medical history make a strong case for a **shared medical decision making setting**, where the patients are expected to take an active role in their health care. Ideally, patients understand the disease and its effects on their bodies, uncertain treatment positive outcomes and the associated negative side-effects.

Patients are typically referred to a rheumatologist by their primary care physician if RA is suspected. Rheumatologists use a number of symptoms and blood tests to determine if RA is present. Radiographs (X-Rays) are routinely ordered by rheumatologists as they can help with the diagnosis, and also to have a record of disease progression.

Radiographs of the hands and feet have been used consistently for the past half century to evaluate RA disease and the degree of progression [48]. The most important criterion for evaluating disease-modifying antirheumatic drugs (DMARDs) in clinical trials has been their capacity to slow down radiographic damage [81]. Magnetic resonance imaging and ultrasound are viable alternatives to plain radiographs and have been shown to be more

Figure 2.2: Example of the pose difference in the PA view for the same set of bones in different patients due to varying RA damage. Left: early stage. Right: late stage.

sensitive in early detection of structural changes [80], however, cost and availability may limit these techniques in practice.

Quantifying joint damage Several methods have been used to evaluate RA disease progression based on radiographs. The first widely used methods are the Steinbrocker and the Kellgren methods [94], where a minimally descriptive global scoring scale from I to IV is given based on the degree of damage to a joint, biased toward the most damaged joint. These methods were not descriptive enough to capture subtle changes due to disease progression. More detailed assessment methods are currently being used, such as those based on the work of Sharp and Larsen [94]. While the two methods are highly correlated [82], the Larsen method is easier to score and less time consuming compared to the Sharp method [94]. The Larsen method [56] includes erosions and joint space narrowing. Each joint is individually assessed with respect to reference radiographs and assigned a score from 0 to 5. The range of the combined score is between 0 to 100 [56].

Radiographs provide ground truth for joint damage, however, subtle progression is difficult to detect with at least a 6 months to 1 year period needed to notice changes in an

individual patient. Modern practice requires treatment prior to radiographic damage [94]. Pincus et al. [83] show that radiographs are strongly correlated with duration of disease and joint deformity. RA patients are typically seen by a clinician due to pain, swelling and functional status. When joint deformity has already occurred, DMARDs and biologic agents are imperative in slowing down progression, but the onset of the beneficial effects may be several months.

2.2 Human Computer Interaction for the Disabled

Disabled individuals can benefit tremendously from computer assistance. Put simply, technology can enhance their engagement with the world. Human-computer interaction (HCI) is a rich area of research with ongoing efforts to build robust and user friendly interfaces. These works target different groups of people with vision, hearing, cognitive and physical impairments, as well as older people. Gallagher et al. [33] are undertaking a critical review of existing technologies for older and disabled people. They reviewed 5143 papers from 2005 to 2012, of which 249 are concerned with older/disabled people. The topics covered in this literature are diverse [33] and fall in categories such as mobility/way-finding, communication and social interaction, access to and use of information, interacting with/using technology, attitudes to/experience of technology, education, using the web, daily life tasks (e.g., memory support) and methods for working with older/disabled people.

Of particular relevance to this dissertation are papers that address interaction with computers. Standard input devices for general purpose computers are the keyboard and mouse. Mastering these devices require fine motor skills and practice. A usability challenge for patients with RA is dexterity, movement restriction due to deformities and pain. Our contribution to HCI is an input method based on hand gestures, easily setup in a doctor's office and highly customizable for a wide range of functional limitations of the hand.

Gestural interaction has become a common means of controlling "intelligent" devices. There are many ways to implement a gesture system, e.g.: whole body movement [12], eye

winks [63], movement of a remote wand [107]. Most of the commonly used gestures may be difficult for patients with dexterity limitations. Our work is similar to Guesgen et al. [41] who proposed gestural control of home appliances by those with mobility limitations.

Our work capitalizes on patients' **gross** motor capabilities through a gestural interface that is comfortable, physically sustainable and easily learned. We designed the system to be a multipurpose input tool: it allows for simple selection operations (e.g., from a menu), while also capable of continuous (i.e., positional) control that can be used to navigate a 3D world.

2.3 Conditional Random Fields

The mathematical tools of probability gives us a way to encode and reason with available knowledge about the world. Often, this knowledge comes in the form of measurements, from which we wish to make *inferences* in light of uncertainty. We call a measurement or an observation a random variable. Random variables can take one of a finite set of values, in which case we refer to them as a discrete random variables or they could take a value from an infinite set, in which case they are continuous. We usually denote random variables with capital letters (e.g.: X or Y).

The probability P of a discrete random variable X being in a particular state x is a number between zero and one:

$$0 \leq P(X = x) \leq 1, \tag{2.1}$$

or short $P(x)$.

The sum of probabilities for all possible outcomes (called the *sample space*) equals one:

$$\sum_i P(X = x_i) = 1. \tag{2.2}$$

We can also quantify the probability of two or more variables sharing a combined state, using a joint distribution:

$$0 \leq P(x_1, x_2, ..., x_n) \leq 1. \tag{2.3}$$

The interaction between random variables can be expressed using *conditional probabilities*:

$$P(x|y) = \frac{P(x, y)}{P(y)} \tag{2.4}$$

$P(x|y)$ is read as the probability of x given y. Two variables are said to be independent if $P(x, y) = P(x)P(y)$.

Declarative representations within general purpose computing devices involve the construction of *models* of *systems* that one wishes to reason about. These representations can be manipulated by various algorithms to answer questions. An important property of declarative representation is the separation between knowledge and reasoning.

Many real-world systems have different processes that are interrelated. We use random variables to encode some knowledge, that, depending on various states of the world, can be random. When we wish to reason about multiple variables simultaneously, we construct joint distributions over the state space, i.e., all possible assignments of some set of random variables. The notation we use for joint distribution is $P(x_1, x_2, ..., x_n)$, which for a set of assignments of x_i is a probability between zero and one.

When the joint distribution is known, we can answer complex questions. For example, we can retain only a subset of the variables in the joint distribution through a process called marginalization. This allows us to ignore (marginalize out) some variables and only reason with the remainder set of random variables. We can also fix a set of random variables to certain values and ask what the *posterior* distribution is, given known values for that set of random variables values.

While working joint distributions offer a powerful tool to answer complex questions, it can become prohibitive to encode the joint state space. One solution is offered by probabilistic graphical models, that, put simply, exploit the structure of complex distributions

by describing them compactly and in a way to make reasoning with them effective. Probabilistic graphical models encode the independencies that hold in the joint distribution. Intuitively, these properties encode knowledge of the form: **X** is independent of **Y** given **Z**.

One type of probabilistic graphical models is the conditional random field (CRF). It is also known as discriminative random field, and, as the name suggests, is a discriminative model. They key idea in CRF models is that in complex distributions we can save resources by **not** modeling explicitly things that we can observe.

In Chapter 5 we formulate a generalized hand skeleton fitting problem for hand radiographs as inference in a CRF. The structure of the CRF allows us to incorporate conditional knowledge from pixel classifiers. Using this flexible framework, we can *infer* latent (unknown) variables (the points of the shape model) using uncertain information from images.

For a comprehensive review of probabilistic graphical models, I refer the reader to [51].

2.4 Machine Learning Preliminaries

At the time of this writing, we have already entered the era of "big data". Information storage is cheap; the promise that machines can *extract information* out of petabytes of data is driving companies and governments to invest more in storage and keep up with the deluge of data. The automatic extraction of useful information, patterns or analysis of data is what **machine learning** can provide.

For example, consider the standard output of a Microsoft Kinect sensor. It provides streams of color and depth information, both at a rate of 30 images per second. For the depth stream, each pixel is typically represented as a 16 bit number (2 bytes). At a resolution of 640×480, the device generates $614,400$ bytes every 30th of a second. Each second roughly 18MB of data are produced. Real-time algorithms are expected to process this amount of data and extract useful information from it. Machine learning algorithms come in three main flavors:

Supervised (or predictive). Essentially, the goal is to learn a mapping from a set of in-

puts x to a set of outputs y, given a set of N labeled input/output **training** pairs $\mathcal{D} = \{x_i, y_i\}_{i=1}^{N}$. Each training input x is a D-dimensional vector of numbers, representing a set of descriptive characteristic of a data point (e.g., weight and height of an individual). The elements of x are called **features** or **attributes**. Similarly, the output values or response variables y can be **categorical**, say an element from a finite set $y_i \in \{1, ..., C\}$ and/or real valued. When the output y is categorical, the problem is known as a **classification** problem. When y is real valued, the problem is known as **regression**.

Unsupervised (or descriptive). In this setting, only inputs are given $\mathcal{D} = \{x_i\}_{i=1}^{N}$, and the goal is to discover *interesting* patterns. By this formulation, the problem is ill-posed since there can be many definitions of interesting.

Reinforcement. The idea here is to use reward or punishment signals to optimize future behavior. Take, for example, the process one undergoes when learning how to ride a bicycle: moving forward is rewarded while falling is punished.

The gesture recognition system proposed in Chapter 4 employs supervised learning, where the training inputs x are histogram based descriptors of a set of hand poses and the outputs are categorical (pose the user's hand is in at a given time).

Parametric vs. Non-Parametric Depending on whether these models have a fixed number of parameters or if the number of parameters grows with number of training samples, they are **parametric** and **non-parametric**, respectively. The gesture recognition model in Chapter 4 is a parametric model, while the model presented in Chapter 5 is a non-parametric model.

Dimensionality Reduction Often, input vectors x contain redundant information when the input dimensionality is high. It is then useful to consider reducing the dimensionality of the input space by a **mapping** or **embedding** to a **latent** space where the latent variables

have desirable properties. For example, it might be useful that variance of the data is maximized in the latent space.

Latent variable models seek a relationship between a D-dimensional input x to a corresponding d dimensional vector t. The most common linear model to express this relationship is *factor analysis* [9]. The linear relationship can be expressed as:

$$x = Wt + \mu + \epsilon \tag{2.5}$$

where W is a $D \times d$ matrix that relates the two variables, while μ allows a non-zero mean model. A common assumption is for the latent variables to be Gaussian with zero mean and unit covariance $t \sim \mathcal{N}(0, I)$. If we specify the noise to be Gaussian $\epsilon \sim \mathcal{N}(0, \Psi)$, this induces a Gaussian distribution on the observations $x \sim \mathcal{N}(\mu, W^T W + \Psi)$. In factor analysis, the error covariance, Ψ, is constrained to be a diagonal matrix. The model parameters can be found using maximum likelihood estimation. Since there is no closed form solution for MLE, one must use an iterative approach [101].

We now explore different choices of the residual covariance matrix Ψ. One possibility is to ignore noise, i.e.: $\Psi = 0 \times I$. This model then becomes:

$$x = Wt + \mu. \tag{2.6}$$

Estimating W has a closed form solution and the method is referred to as principal component analysis (PCA). PCA is a linear model that finds W, such that the data reconstruction error is minimized. Consider a data matrix $X = [x_1, x_2, ..., x_N]$ of N column vectors $x_i \in \mathcal{R}^D$. We seek an orthogonal matrix W to project points in X onto $\mathcal{R}^d$ where $d << D$. This problem can be formulated as a minimization of the reconstruction residuals:

$$\arg\min_{W} ||X - WZ^T||_F^2 \tag{2.7}$$

subject to the constraint that W is orthonormal. Z is an $n \times d$ matrix with z_i in its rows and $||A||_F$ is the Frobenius norm defined as:

$$||A||_F = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2} = \sqrt{Tr(A^T A)}$$

It can be shown that the optimal W is $\hat{W} = V_d$ where V_d contains the d eigenvectors with largest eigenvalues of the empirical covariance matrix $\hat{\Sigma} = \frac{1}{N}\sum_{i=1}^{N} x_i x_i^T$ and the optimal latent representation of the data is $z_i = W^T x_i$. This is the standard formulation for PCA.

Tipping et al. [101] formulate a probabilistic version of PCA, where the assumption on the model noise is isotropic Gaussian $\epsilon \sim \mathcal{N}(0, \sigma^2 I)$:

$$x = Wt + \mu + \epsilon \tag{2.8}$$

The t-conditional distribution over x is then given by:

$$p(x|t) \sim \mathcal{N}(Wt + \mu, \sigma^2 I) \tag{2.9}$$

If we assume $x \sim \mathcal{N}(0, I)$ then we can integrate latent variables and show that [101]

$$t \sim \mathcal{N}(\mu, C) \tag{2.10}$$

where $C = WW^T + \sigma^2 I$, with additive noise σ, and $\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$ is the sample mean. Similar to PCA, there is an analytic solution for W and σ using the eigendecomposition of the $WW^T = UL^2U^T$:

$$\sigma^2 = \frac{1}{D-d}\sum_{j=d+1}^{D} L_{jj}^2 \tag{2.11}$$

and

$$W = U_d(L_D^2 - \sigma^2)^{\frac{1}{2}} \tag{2.12}$$

We denote U_d as the truncated version of U retaining only the first d columns, and L_d^2 is a truncated version of L^2 retaining only the first d columns and rows. Intuitively, σ^2 is the variation discarded by the components ignored over the lost dimensions.

Kernelization When solving for W in the PCA and PPCA models, we used the inner product $x^T x$ as a notion of similarity. If X contains points that are not linearly separable in N dimensions, we may wish to map the points into a higher dimensional space, where they can be linearly separated. Suppose we have a function ϕ that maps an input vector $x \in R^D$

to a higher dimensional space $\mathcal{F}$, i.e.: $\phi : \mathcal{R}^D \mapsto \mathcal{F}$. The inner product $\phi(x)^T\phi(x')$ may be computationally expensive to get. By Mercer's Theorem, it is enough to equip the input space $\mathcal{R}^D$ with a measure of similarity and show that, for a class of positive semi-definite **kernel** functions $k(x, x') \in \mathcal{R}$, $k(x, x') = \langle\phi(x), \phi(x')\rangle_{\mathcal{F}}$ is an inner product in $\mathcal{F}$. The key insight here is that we never have to explicitly represent ϕ. The **kernel trick** involves computing the inner product matrix in $\mathcal{F}$ via a kernel function that satisfies Mercer's condition. Below are some examples of commonly used kernels:

- Linear

$$k(x_i, x_j) = x_i^T x_j \tag{2.13}$$

- Degree p polynomial

$$k(x_i, x_j) = (x_i^T x_j + 1)^p \tag{2.14}$$

- Radial basis function

$$k(x_i, x_j) = e^{-\frac{1}{2}(\frac{(x_i - x_j)^T(x_i - x_j)}{\lambda^2}} \tag{2.15}$$

Chapter 3 Decision Aids for RA Patients

Patient decision aids (DAs) are becoming a critical part of clinical care. They are crucial, given the limits on the time practitioners can spend with patients. They serve as educational material and help guide patients' questions. In this chapter, we survey existing DAs for rheumatoid arthritis (RA) patients considering a treatment plan. Our findings indicate that the few DAs available to RA patients are lacking in two key areas: effective communication of probabilities and citations of evidence. We also address factors that impact chronic disease patients decision-making skills, such as health literacy and numeracy, and willingness to be an active participant in shared decision making. We suggest that developers of DAs follow recognized methods and guidelines such as those in the International Patient Decision Aid Standards (IPDAS) [1] during the development process. We also emphasize the importance of customizability and interactivity of DAs in order to make the decision process more visceral, while ensuring the avoidance of presentational bias and taking into account the unique values and preferences of patients.

3.1 Introduction

Rheumatoid arthritis (RA) is a potentially debilitating, life-long autoimmune disease that affects millions of adults around the world. Individuals suffering from RA face a daunting array of treatment choices, each with its own benefits and side effects. The challenge in making such choices is often difficulty in comprehending the consequences of a particular choice of medications and other treatments. The choice patients make is potentially life-altering and overwhelming. Enhancing patient knowledge of disease outcomes and treatment options is beneficial, because it can greatly reduce patients feelings of uncertainty and increase their confidence in the decision making process [86, 75].

Communication of treatment benefits and potential complications of disease and treat-

ment is essential to medical care. Patients with RA have multiple options from which to make complex decisions. Each drug offers potential benefits and risk profiles that patients may value differently [40]. For example, biologic agents may induce remission, but may incur significant out-of-pocket costs and risk of serious infection or malignancy. When patients participate in shared decision making, they want information about alternatives and the ability to assess risk and ask, "*What if?*"

Aside from the emotional trauma when informed of the diagnosis, patients face the non-trivial decision of accepting a treatment. Good, shared health care decision-making requires a quality patient-clinician interaction, which is often hindered by time constraints. To partially overcome this, patients receive verbal explanations, often combined with static decision aids (DAs) (e.g., printed materials) [75, 28]. They also seek information from other patients and search the Internet to find relevant information that may not be scientifically sound. Patients are expected to understand their disease, treatment options, and associated risks, and to be competent to partake in treatment decisions. However, patients with low health literacy may not fully understand treatment options, risks and benefits from written material. Conversely, some patients may be highly literate and understand the material, but some existing DAs fail to provide citations [29] that health literate patients may further pursue. Another shortcoming of current practice is that the written and verbal explanations may frame information such that the final decision is biased towards a specific outcome [26, 29]. This chapter addresses the current practice in communicating uncertainties about medications to RA patients using DAs. We also look at important factors that make DAs for RA successful and suggest improvements, such as dynamicity or interactivity of a DA using a computer or gaming console.

3.2 Probabilities and Risk

Many studies have shown that the use of written materials and other static tools to present treatment option information about RA and educate patients is ineffective for many reasons,

the most significant of which is low health literacy and numeracy [88, 35, 73, 102, 10]. Health literacy is defined as "the degree to which individuals can obtain, process, and understand the basic health information and services they need to make appropriate health decisions" [73]. Numeracy refers to the degree of one's competency to use numerical information in one or a few short calculations in order to solve a problem [35]. Walker et al. [114] explored the relationship between health literacy and knowledge gain from an arthritis information booklet, accompanied by a "Mind Map" (dramatic words and images to aid cognitive processing). One group received the booklet, while the other received the booklet and the "Mind Map". Both groups gained some knowledge, but there was no evidence of the "Mind Map" helping, regardless of health literacy assessment. Such evidence suggests that we need a different approach to communicate uncertainty and information about treatment effectiveness from clinical trials.

It is generally difficult to process and understand probabilities [34]. Gigerenzer suggests that human evolution led to the development of cognitive inference machinery in order to adapt to risky situations; however, the format of the information that we use naturally does not come in probabilities or percentages, but absolute frequencies [34]. As an example, it is often easier for people to understand that during a clinical trial, 2 out of 1000 patients experienced severe side effects in contrast to 0.2%. Inferences are reasoning processes on the basis of sometimes incomplete, circumstantial evidence and prior conclusions. Patients and health care practitioners have to make inferences based on numerical data from evidence-based clinical trials. These results reveal a need for clinicians to pay closer attention to the educational materials they distribute to patients. Problems can arise when the numerical information provided require the computation of inferences where additional evidence is learned and probabilities of events are updated due to new knowledge. The new information that updates existing probabilities should be incorporated through the use of Bayes rule, something that many patients are unable to do. Bayes rule provides a way to update probabilities based on existing evidence.

In summary, RA patients are faced with decisions for which they often lack information and understanding of the options. The use of DAs can, potentially, help patients make more informed decisions, and thus increase adherence to those decisions. A good DA should convey the possible outcomes, both positive and negative, of treatment, and their likelihoods, in ways that patients can understand. It should enable patients to learn more, perhaps through the use of "What if?" scenarios or other interactions.

3.3 RA Treatment Options

The clinician's goals in treating RA are remission, improved function, and prevention of deformity. The patient's goals include pain relief and improved function. Achieving these goals with minimal to no risk is desired but often impossible. An American College of Rheumatology (ACR) committee of experts used a formal group process to review scientific evidence to create recommendations for use of nonbiologic disease modifying antirheumatic drugs (DMARDs) and biologic agents in RA [89]. They emphasize that these recommendations are "intended to guide therapy" [89] in providing personalized patient care. No randomized controlled trial (RCT) evidence exists to support preferential use of one monotherapy over another. There are over 170 possible dual- and triple-DMARD combinations among five nonbiologic drugs [89]. In addition, three classes of biologic drugs (five tumor necrosis factor inhibitors, abatacept, and rituximab) can be used, usually in combination with a nonbiologic DMARD. A new class of biologic agent, tocilizumab, is excluded because it received FDA approval after publication. More biologic agents are in the pipeline. Thus, the decision to choose a treatment plan for RA is incredibly complex and doesn't fit a simple clinical guideline. In this chapter we focus on surveying existing DAs for patients with RA.

3.4 Difficulties in Choosing a Treatment

Patient participation is particularly relevant for many rheumatic diseases [23], compared to other conditions in which a treatment decision is urgent. In situations in which there is no time pressure associated with selecting a treatment, the patient can be given information to digest at home and potentially become an active participant in shared decision making where the trade-offs between treatments are closely knit [23]. Ideally, patients comprehend their role in the process as seen by physicians; however, this is not always the case. Patients can be divided roughly into two groups based on their interaction style with the physician [16]: those who are more passive and accept a paternalistic approach by a clinician, and those who are more autonomous because they understand they are active participants. Haugli et al. [43] found that patients with RA wished to be seen holistically by their physicians. Thus, Interaction style becomes a factor in the decision making process [43]. RA affects every aspect of patients lives due to consistent pain, inability to perform routine tasks and deformities that develop in later stages of the disease. The feeling of being understood by physicians provides patients with a sense of security and emotional support during times of hardship and vulnerability [16].

Hirono et al. [45] address the relationship between participation style and the feeling of being understood by the physician. They conducted a qualitative study and found that patients who perceived themselves as having actively participated in the visit felt they were better understood by the physician. Conversely, those who were less active in the decision making process during the visit felt less understood. This suggests that if we make the assumption that patients will participate after receiving any type of information, it may not enhance their feeling of being understood. Patient involvement is directly correlated to decision satisfaction. O'Connor et al. [77] studied the impact that DAs have on risk communication prior to interaction with practitioners. They determined that, if a DA was used, the quality of the time spent with the provider was better, and decision satisfaction was increased. Thus, patient-clinician style has been shown to affect patient decision making.

In addition, DAs can help patients, whether or not the patients are otherwise involved in their own healthcare decision making.

3.5 Patient DAs

Todays patients are faced with difficult decisions to choose one treatment from several available options with probabilistic outcomes. Their reasoning is biased by emotion [92], difficulty processing numerical data and misconceptions after reading literature that they consider pertinent to their condition [92].

DAs are tools to increase patients' knowledge of options and facilitate their involvement in the health decision-making process, while taking into consideration cognitive biases, possible information overload, vocabulary and avoidant coping [84]. O'Connor et al. [76] state that DAs are an invaluable addition to usual clinical care. However, DAs often present incomplete and uncertain information, so care must be taken to create DAs that reduce uncertainty without increasing patient anxiety [84].

Evaluating DAs is a complex process. Most trials focus on the short-term impact of the decision and knowledge enhancements they provide [84]. McCaffery et al. [84] suggest a new approach to evaluating DAs; they focus on the long-term effects in terms of patients' quality of life. They argue that while a DA might make the decision process longer and more complex, if long term quality of life is improved, then it can be considered successful as opposed to one that makes the decision process quicker, simpler and more rewarding at the time when the decision is made, but leads to a decrease in quality of life. Thus, DAs should be designed with long term effects in mind, with the goal of minimizing decisional conflict and maximizing quality of life in the short, medium and long term.

Schwab [92] claims that patients or their families are unlikely to have a fundamentally different mental construct of the decision than the physician or practitioner has. The disagreements appear when the decision is made about which option should be chosen [92]. Here, a DA can arguably act as a moderator, biased by the patient's preferences, and max-

imize expected utility for the patient. A fundamental concern for any DA is potentially intractable demands on decision-makers, leading to general misunderstandings conducive to uninformed decisions [92]. This may result in decreased longevity and reduced quality of life for the patient.

In short, DAs can be extremely helpful, but a DAs effect on decision making is only as good as the DA itself. Quality of DAs can be measured in terms of completeness of information, clarity and correctness of presentations, the ability to personalize the DA with the individual patients condition and preferences, the weighting of short and long term quality-of-life issues, and the appropriate use of probabilities in computing expected outcomes.

3.6 Available DAs for RA Treatments

The diagnosis of RA is life altering. Patients consult clinicians, the Internet, friends, and family for information. The rheumatologist's role is to help patients consider disease duration and severity, comorbid conditions and poor prognostic factors to guide therapy. The ACR 2008 RA treatment recommendations are intended to promote beneficial outcomes and permit individualized treatment decisions [89]. However, complicating factors exist.

1. Few providers are sufficiently trained in effective patient communication, especially about probabilistic information. Most statistics or probability classes during providers' training are not designed to foster insight for patients seeking more information. For example, as Gigerenzer suggests, "for every confusing representation there is at least one alternative, such as natural frequency statements, which always specify a reference class and therefore avoid confusion, fostering insight." [35] Communication strategies include printed information and counseling. DAs support decision-making and supplement patient-clinician interaction. In the University of Kentucky (UK) Rheumatology Clinic, patients routinely receive ACR Patient Fact Sheets [2] to supplement unstructured discussions. However, health care economics limit face-to-face

time, and fact sheets do not present probabilistic information or clarify patients' values [32, 102].

2. Trust in physicians has a significantly greater effect on patients' confidence in decision-making than specific knowledge of drugs and diseases [11]. Trust develops during a long-term relationship, but patients seen in consultation are often asked to make treatment decisions at the first visit. Thus, they learn about RA and treatment alternatives when trust is barely developed. Explanations for unwarranted variations in clinical practice include under-use of effective care, and variations in preference-sensitive care and/or supply-sensitive care [105]. Providing patients with information about their disease and treatment choices leads to shared decision-making. This changes the pattern of preference-sensitive care, which is usually dominated by medical opinion [86, 102, 105].

DAs can increase understanding, but are more effective if structured, tailored and/or interactive [114, 91]. Walker et al. [114] have shown that the ARC booklet increases knowledge in functionally literate patients and the mind maps had no effect. This was contrary to expectations, since the mind map was intended to aid cognition visually through diagrams and images, thus suggesting different approaches need to be considered. Visual graphics that display risk information can aid understanding and supplement counseling by providing information about options and outcomes and by clarifying personal values related to benefits and harms [91]. For example, pictographs can be used to represent probabilities in a format that allows one to count icons in a grid, where different colors or versions of the icon represent a probability class and the total number of icons is known. Using pictographs limits biases based on anecdotal information from other patients, effectively communicates medication side effects, and reduces side effect aversion in decision-making [91, 115]. To date, interactive, accessible DAs for RA remain undeveloped. In the following subsections, we look at specific treatments and the DAs that are available for those treatments.

3.6.1 Methotrexate

Currently, methotrexate (MTX) is the most widely used first-line treatment for RA, and is also used as the backbone in combination with newer biologic agents [68]. Initially introduced in oncology, MTX was found to be effective in RA despite a lack of understanding of its mechanism of action. It is proposed that MTX reduces toxic metabolic compounds in chronically inflamed tissues [72]. MTX acts as an immune system suppressant, which may be beneficial for controlling autoimmune-mediated inflammatory diseases; however, it may potentially increase the risk of infection. The medical literature contains conflicting evidence regarding a higher risk of infection due to treatment of RA with low dose MTX [68]. Practitioners must convey these uncertainties to patients who may have an ill-conceived notion that treatments are 100% effective. Therefore, when conditioned on their own values and preferences, patients may consider not receiving any treatment. Refusal of treatment may be a suboptimal choice in light of existing medical evidence, but the patient may perceive it as optimal due to natural risk aversion tendencies. It is thus imperative that a good DA allows for such an outcome.

OConnor et al. [103] developed a patient DA for patients that contemplate taking MTX. This DA starts by explaining the basics of how RA is a disease of the immune system and attempts to clarify the uncertain nature of both the clinical outcomes of MTX and also the uncertain quality of the studies from which they retrieve the information contained therein. The authors of the DA mention the information in the DA comes from 7 reviews, out of which 2 are Cochrane reviews plus 8 more studies; citations for the studies are not provided. The DA authors frame study results in terms of improvement rate (e.g. ACR 20 response: the number of people out of 100 that showed 20% improvement with MTX and a placebo). These numbers are also given in the form of pictographs. They also have several paragraphs comparing MTX to other DMARDs, concluding that MTX is the best choice as the first line of treatment in RA.

Alongside the informational material, the DA developers provide a worksheet with 6

steps for patients to help structure their thoughts and feelings. This worksheets purpose is to be shared with their clinician in order to help them communicate more effectively. In step 1, the patients are asked to think about when the decision has to be made. The purpose of this is to help the patients think about the urgency of treatment selection in light of uncertainty given existing evidence about MTX and their own values and beliefs. Step 2 asks patients to consider the pros and cons of MTX by eliciting preferences regarding each pro/con by circling a number of stars out of 5, 5 stars being considered most important. Step 3 is intended to elicit the patients self-perceived role in the decision making process (e.g., preference that clinician makes the decision for them or their own involvement while asking opinions of others.) Step 4 is intended to elicit extra information or help that the patient may need to make the treatment decision. Step 5 elicits more information in the form of an open question regarding their perception of the next steps involved in the decision making process. Step 6 asks the patients to share the information in the worksheet with their clinician.

Rader et al. [85] created a static DA to help patients decide between taking MTX alone, MTX in combination with other DMARDs and further discussing the problem with their practitioner. Their DA has 4 steps:

1. Benefits and side effects of each option.
2. Preference elicitation to help patients think about what matters most to them.
3. A quiz to determine information gain.
4. Next steps (decision to take MTX alone, MTX with other DMARDS or further discussion with practitioner)

The patient is first given basic information about RA and MTX, and then shown the current risks and benefits from the latest evidence-based trials comparing MTX alone or in combination with other DMARDs. This is accomplished via four pictographs, which were

shown to be effective in risk communication [115]. The quality of the evidence is classified into four categories, as follows [85]:

- + Very low quality – further research is very likely to change the estimate.
- ++ Low quality – further research is likely to change the estimate.
- +++ Moderate quality – further research may change the estimate.
- ++++ High quality – further research is very unlikely to change the estimate.

The benefits of MTX treatment in combination with other drugs are marked with ++ (low quality requiring more research) and side effects with +++ (moderate, requiring more research) based on the quality of the medical trials. The explanations above are given as a footnote, which may be overlooked by some patients. The + signs may also be confused with quantity. Evidence is stronger for benefits (marked with ++) than for side effects (+++), which may be incorrectly interpreted as more side effects than benefits. They provide a reference for the source of the statistics, which may be beneficial for those who want more information. This DA does not have an option for no treatment, which we regard as problematic.

The developers of this DA evaluated it with respect to IPDAS (International Patient Decision Aid Standards). IPDAS evaluates DAs in three major areas: content, development process and effectiveness. Rader et al.s MTX DA scored 14 out of 15 in the content area, 6 out of 9 in the development process area and 0 out of 2 on the effectiveness criteria.

3.6.2 Biologic Agents

The Ottawa Hospital Research Institute (OHRI) has made available to the public DAs for biologic agents used in RA treatment. All DAs for RA follow the same format as Rader et al.s DA for MTX in combination with other DAs. These DAs contain a paragraph about what RA is and a paragraph describing the medication (abatacept, etanercept, tocilizumab)

and how it works. The benefits and side-effects are rated based on the quality of the evidence from clinical trials with the same (+) notation as described above. Thus, these DAs would have the same problems as that for MTX. In addition, they communicate results about benefits and side-effects using pictographs. Chilton et al. [20] explored RA patient treatment preferences and decision-making when faced with three anti-TNF-alpha inhibitors (etanercept, adalimumab and infliximab). The study was done on the basis that, when treatment options with similar clinical outcomes had different effects on quality of life, the patient should be offered the choice. Adalimumab and etanercept are administered by subcutaneous injection, which may be performed safely by the patient or family members, while infliximab is administered intravenously by a trained health care professional. The methods of administration have a direct impact on the patient's quality of life. In their study, 190 participants were given a questionnaire to complete via mail and seven were interviewed in person. The results provide insight into a wide range of patient decision making processes regarding treatment selection, which are critical in the design of a DA. For example, patients require reassurance and support from their health care practitioners When this is combined with a better understanding of risks and trade-offs, patients are more likely to participate in the decision process. Even a good DA cannot replace time with the practitioner, but it can have a large impact on the patient's understanding of the treatments. Preferences vary among patients. Some learn how to perform the subcutaneous injection rather than travel to the infusion center, while others are anxious about self-administration and feel more comfortable when they are in contact regularly with health care professionals and other RA patients. The most preferred choice of treatment was adalimumab due to ease of administration and its availability as a pre-filled injectable medicine.

3.7 Conclusions

Rheumatoid arthritis is a chronic disease with a relatively slow progression rate. This puts RA patients at a slight advantage over sufferers of most other diseases due to the lack

of urgency to select a treatment that is in tune with their values and preferences. Such a patient-centered decision process takes time, and the decision can have long lasting impact on the patient's long-term quality of life. It is therefore important for providers to make available a quality DA for patients, complementary to traditional patient-clinician interaction. Few RA DAs are available at the time of this writing. Those that do exist are lacking in key areas of effective probability communication and evidence citing. As noted above, DAs are more effective if structured, tailored and/or interactive [114, 91]. In addition, our review of existing DAs for RA lack personalization that would be useful in eliciting preferences and values unique to every patient. We conclude that developers of DAs need to follow recognized methods, cite valid sources and avoid presentational bias, while ensuring the unique values and preferences of patients are taken into consideration.

Chapter 4 Kinect Based Hand Gesture Recognition

This chapter has appeared in the printed proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV) 2012 [71]. In this chapter, we propose a robust static hand gesture recognition algorithm that makes use of two Kinect sensors. This can be used to control an avatar in a decision aid for rheumatoid arthritis patients who have a difficult time using a standard keyboard and mouse interface. The sensors are placed on the left and right sides of a target sensing area, easily set up in a doctor's office or waiting room. The Kinects provide a rich point cloud, out of which gestures from a known vocabulary are recognized in real time. We use 6 point cloud descriptors simultaneously and employ the majority rule voting scheme to pick a "winner" gesture in real time. We achieve rotation invariance by using part of the forearm as a good indicator of hand orientation and aligning the hand with the world coordinate system origin. We evaluate the performance of the recognition system under various motion and rotation conditions.

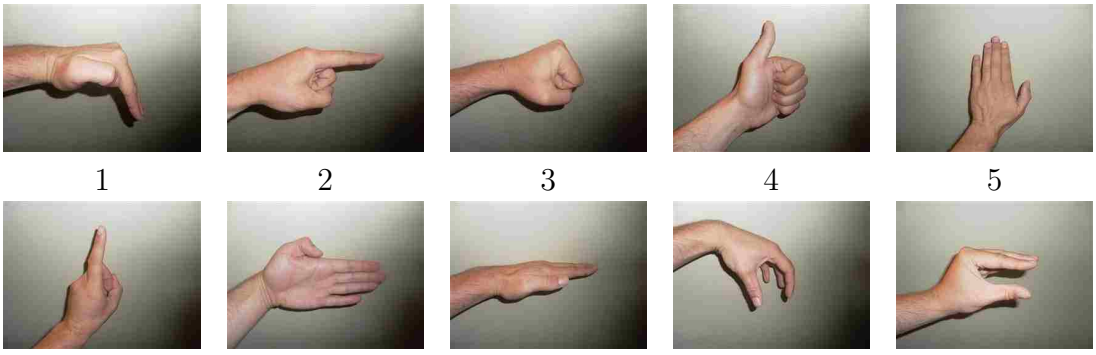


Figure 4.1: Gesture vocabulary.

Rheumatoid arthritis (RA) is a potentially debilitating, life-long autoimmune disease that affects millions of adults around the world. Individuals suffering from RA face a daunting array of treatment choices, each with its own benefits and side effects. The challenge in

making such choices is often difficulty in comprehending the consequences of a particular choice of medications and other treatments. To the best of our knowledge, decision tools for RA based on computer game technology have not been explored. The problem for RA patients is that the standard modes of interaction are often impossible because the disease limits range of motion and can cause physical deformities: simply manipulating a mouse or grabbing a Wii-mote may be impossible.

In this chapter, we propose a real-time gesture recognition system that uses a pair of Kinect sensors to distinguish between static hand gestures. Our system is designed to be easily customized for individual users and to be rapidly configured in a doctor's office. Given the physical deformities and limited range of motion, each user may require a special set of gestures that are determined interactively by a therapist. Therefore, we minimize training time by using a lazy-learning algorithm to classify individual gestures.

In the context of our larger research program, this gesture recognition system will enable RA sufferers to control a character in a virtual environment. The goal is for them to use this system to visualize the outcome of a particular treatment plan by enabling them to perform actions, such as making a cup of coffee, which might be impossible for them in the physical world. For this task, our gesture recognition system might include gestures for grabbing, standing, walking, and placing-on-a-table to control avatar motions. Figure 4.1 shows examples from the library of natural static hand gestures we use in our experiments. The heart of our proposed system is a voting-based scheme that combines the results of 6 nearest-neighbor classifiers to determine the current gesture. The resulting system works in real time on a modest computer with average resources. An important feature of our method is that it does not require tracking or manual pose initialization; these techniques were deemed too brittle for our proposed users, and, as our results show, were not necessary. Instead, we classify individual gestures on a per-frame basis using a single point cloud, fused from two Kinect sensors, as input. The most computationally demanding part of our system is the component that manipulates the point cloud, namely a rigid-body

transformation and a hand segmentation procedure. Once segmented, we convert the point cloud into a voxel-based representation and extract 6 different features.

Our results demonstrate that the proposed system can accurately classify a realistic set of gestures in real time. We demonstrate that the system can be rapidly retrained for new users. Future work will more extensively evaluate the system and integrate this gesture recognition system into a larger virtual environment to aid RA sufferers to make informed medical decisions.

4.0.1 Related Work

Recovering the full kinematic parameters of the skeleton of the hand over time, commonly known as the hand-pose estimation problem, is challenging for many reasons: high dimensionality of the state space, self occlusions, insufficient computational resources, uncontrolled environments, rapid hand motion and noise in the sensing device [27]. Erol et al. [27] provide a comprehensive review of research on this problem. Given our application, we focus on a special case of hand-pose estimation known as gesture recognition, in which discrete hand poses are detected. Solutions to this problem can be generally divided into appearance-based methods, depth-camera methods, and tracking-based methods. In the remainder of this section we give an overview of the gesture recognition literature.

Direct Appearance-Based Methods Athitsos et al. [7] propose a real-time gesture recognition system that uses a large database of synthetically generated images of hands in various configurations. The proposed approach relies on searching a database of tens of thousands of potential gestures. They use an indexing scheme, known as BoostMap, to decrease query processing time and thus enable real-time recognition. For their system, the input data consists of 2D images, out of which the hand silhouettes are extracted through skin color segmentation. Their method relies heavily on a relatively clean segmentation. Unlike typical appearance-based methods, where estimation is done from a limited num-

ber of viewpoints such as Ying et al. [108], Athitsos et al. [7] allows arbitrary views. [42] Hassan et al. propose a method using 2D images from a simple camera using the hand contour information and complex moments which are rotationally invariant. The complex moments and contour information are used as input to a feedforward neural network that classifies a vocabulary consisting of 6 gestures. They obtain accuracy of 86.38%, which is lower but comparable to our approach.

Direct Depth-Based Methods Gesture recognition has been revisited with the introduction of inexpensive depth cameras. Most similar to the current chapter, is work by Suryanarayan et al. [97] in which they propose a gesture recognition system using a single ZCam camera from 3DV Systems. One limitation of their approach is their method of obtain invariance to hand rotation. Their method introduces a limitation on the types of poses that can be successfully distinguished. Our hand/forearm segmentation method, coupled with a similar PCA-based normalization scheme, enables our system to accurately distinguish between a more varied set of gestures. Zhou et al. [87] proposed a method to recognize hand gestures using a Kinect sensor. Their approach uses color, as well as depth information. The color is used to segment the hand from the rest of the environment, while the depth is used in a template matching algorithm where the dissimilarity measure they use is "Finger-Earth Mover's Distance" (FEMD). They claim an average of 90.6% accuracy on a dataset, but mention nothing about distortion or rotation invariance.

Tracking-Based Methods An alternative approach to gesture recognition involves first solving the hand-pose estimation problem. Once this is solved it is straightforward to determine the gesture based on geometric parameters of the hand. Hand-pose estimation methods often rely on tracking, where an initialization step has to be performed, and the pose at the current frame relies on knowledge about the previous frame [59, 109]. In such applications, if the system loses track, it can only be recovered using manual initialization. Unlike these systems, our approach estimates pose from a known vocabulary at every

frame, independent of previous frames.

4.1 A Real-Time Gesture Recognition System

We propose a system that uses point clouds obtained from two consumer depth cameras to recognize gestures in real time. In this section, we first describe the intended use of the system and then give details of the physical and conceptual configuration of our system.

4.1.1 Usage Scenario

We designed the system to be robust and easy to set up in a doctor's office environment. The two Kinect sensors can easily be placed on a desk, and the system runs on an inexpensive machine. Our choices of gestures are familiar gestures that will be useful for our application. We designed this algorithm for natural interaction and control of an avatar through gestures. A simulation involves directing the avatar through hand gestures to perform certain activities.

Imagine that the user wishes her avatar to go to the kitchen, get a cup of coffee, come back, and put the coffee on a table. This may sound trivial to the reader, but for someone with advanced RA, this may be impossible. Standing up and sitting down are painful, sometimes impossible. Holding the coffee cup in one hand requires a steady grip; we have heard patients describe walking across the room with a cup, and suddenly the cup and coffee are on the floor.

We want to show the user an avatar that attempts this sequence of actions. At one point in the process, the avatar will display the user's degree of RA and associated deformation. At another point, the avatar will display a possible effect of the medication (calculated according to the probabilities of effects in the medical literature). The user, however, will not suddenly get better in the doctor's waiting room. So control of the avatar must be doable, even with RA-inhibited or deformed hands.

The user will be able to gesture with a finger pointed up to stand, a horizontal palm

to turn, etc. These gestures will all be user-tested for comfort and performability. We are preparing an IRB application for a further study.

4.1.2 Data Acquisition

We create a point cloud using two Kinect sensors. In this work, we ignore color information. There are two main reasons why we can ignore color: first, the hand is relatively flat textured and second, we can not assume consistent lighting conditions. In fact, our system works in complete darkness. While ignoring color, we acknowledge that it could provide additional useful information for solving the hand pose estimation problem. In our experiments, the yaw angle for which the device has a motor is set at the default value, which results in an orientation parallel to the table top on which the sensors are placed. An angle of 45° provides the most information when performing gestures that have inevitable occlusion (e.g., gesture 5 from Figure 4.1), because within a certain distance from the cameras, each of the two sensors "sees" half of the hand, hence capturing more structural information.

We map raw depth values for each Kinect, which are 12 bit integers, to metric 3-space using Nicholas Burrus' formula [14]. Because the two video streams come from two separate sensors, we have to perform a calibration step, which rotates and translates the coordinate system of one camera to match the other.

4.1.3 World Coordinate System

The rotation around the Y-axis previously determined in the calibration step (Θ) is part of the transformation that maps one Kinect's coordinate system to the other. We create a world coordinate system, where the depth axis points exactly in between the two Kinect sensors. This new depth axis coincides with the bisector line of the normal vectors with the origins at the center of the camera sensors as depicted in Figure 4.2 by Zworld. Angle Θ is known from the calibration step, so in practice we rotate the point clouds around the world

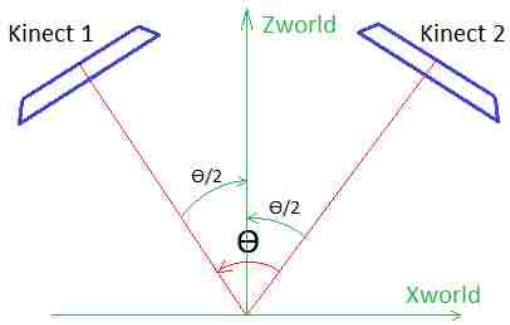


Figure 4.2: World Coordinate System depth axis.

coordinate system Y-axis by $\frac{\Theta}{2}$ and $-\frac{\Theta}{2}$ (clockwise and counterclockwise respectively). We create this coordinate system because a user will interact with the system by placing their hand between the two sensors with the arm pointing along our world coordinate system depth Z-axis. This step is useful in creating a descriptor (Section 4.2.3) where the real depth coincides with the line from wrist to the tip of the fingers. Having an intuitive world coordinate system is invaluable if we need to extract the spatial position and orientation of the user's hand when performing a certain gesture. In our proposed application, a user may navigate the virtual 3D world by pointing with their index finger to determine what direction the avatar will walk next. When recognized, this gesture will cause the avatar to turn her head in the direction where the user's finger is pointing. After a direction has been established, another gesture can be used to make the avatar walk forward or back up. If, for example, the user wishes to pick up an object in the virtual world, the relative position of the hand in our world coordinate system can be translated into that of the virtual world so the avatar can move her hand in a 3D position indicated by the user. The relative position of the hand is the center of mass of the point cloud derived from the segmentation process.

4.2 Gesture Recognition

We formulate the gesture recognition problem in the context of this chapter as follows: given a snapshot of a hand configuration (a point cloud), decide which gesture from a

repertoire of known gestures has been performed. Using a cluttered point cloud requires a segmentation process, to isolate the hand from the rest of the scene. We use a simple segmentation process described in the next section. The simplicity of the segmentation is acceptable due to the way we engineered the system to allow for the hand to be the closest object to the Kinect sensors.

4.2.1 Segmentation

Given the final point cloud Γ=$\{p_1, p_2, ...p_n\}$ in the world coordinate system mentioned above, we need to extract the points that belong to the user's hand. We require that, during use, the hand will be the closest object to the two Kinect sensors. We extract closest point $p_{closest} = (X, Y, Z)$. We then search Γ for a subset of points $\Gamma' = (hp_1, hp_2, ..., hp_n)$, where the following conditions hold.

1. $p_{closest}.Z < hp_n.Z < p_{closest}.Z + 0.30m$
2. $p_{closest}.X - 0.20m < hp_n.X < p_{closest}.X + 0.20m$
3. $p_{closest}.Y - 0.20m < hp_n.Y < p_{closest}.Y + 0.20m$

The subset Γ' is guaranteed to be contained in a box with volume $0.30m * 0.20m * 0.20m = 0.048m^3$. We chose 0.30m for the depth because it captures part of the forearm, which we use to achieve rotation invariance, process described in section 4.2.2. The rotations with which we wish to achieve invariance to are about the X-axis and Y-axis of the world coordinate system. The values 0.20m for width and height of the bounding box, were determined empirically to ensure it can contain hands of various sizes. Assuming the segmentation process completes successfully, Γ' contains points pertaining to the hand and part of the forearm.

4.2.2 Rotation Invariance

To achieve rotation invariance, we assume the forearm is a strong indicator of the hand orientation. We find the principal component of this set and align Γ' along the X and Y components of the world coordinate system. It is true that gestures such as "Palm Front" and "Attention" will affect the principal component vector, and thus final rotation, but the bias is consistent across users, so it can be ignored. We call the angle that the principal component is rotated along the Y-axis of the world coordinate system α and the rotation angle along the X axis, β. This results in invariance to the two rotations. Unconstrained rotation along the Z axis is important because the semantics of the gestures depends on it, e.g., gestures 7 and 8. The user sits in a chair and uses one hand to interact with our system, thus scale invariance is not a concern in our application.

4.2.3 Descriptors

To discriminate between gestures that users perform, we require a fast classification algorithm to process every frame of the combined streams. Therefore, designing a system that is computationally feasible on an average machine is important due to a requirement of cost efficiency to implement and deploy systems in various settings. In Section 4.2.1 we describe the segmentation process used to isolate hand and forearm points from the rest of the background and align it with the world coordinate system to achieve rotation invariance, and we named this set Γ'. The forearm points were used for alignment purposes, and can now be discarded. We eliminate them by "trimming" Γ' of the trailing 15cm volume.

In order to build a gesture vocabulary and implement a classification algorithm, we need to compute a descriptor for Γ'. We use six point distribution histograms for Γ'. We subdivide the volume into $6^3, 8^3$ and 10^3 voxel spaces of equal size, distributed evenly.

For each such division, we produce two descriptors; one where each dimension of the descriptor corresponds to the point count in each voxel space and the other's dimensions are binary values describing whether there is one or more pixels in each voxel space. We

use Ψ to denote a count histogram and ψ to denote a binary histogram. The final descriptors are $\Psi_{6^3}, \psi_{6^3}, \Psi_{8^3}, \psi_{8^3}$ and Ψ_{10^3}, ψ_{10^3}. Since the hand is extracted from the world coordinate system, the position is important to our application. We determine the absolute position of the hand by computing the center of mass for the point cloud Γ' and name it Ω – a vector in 3-space.

4.2.4 Recognition

We use a nearest neighbor classifier in combination with majority rule voting scheme to recognize unknown gestures. For each unknown gesture, we compute the six descriptors and map the points to the feature spaces of each descriptor. We compute distances between all cluster centers. The recognition process for an unknown gesture involves computing the squared distance in each feature space between it and known gestures in the vocabulary,

$$d^2(\Psi_{unknown}, \Psi_{known}) = \sum_{n=1}^{i}(\Psi_{unkown}[i] - \Psi_{known}[i])^2$$

where i is the dimensionality ($6^3, 8^3$ or 10^3).

We retrieve the top two matches, and the threshold that determines whether the gesture is known or not is twice the distance between the top two cluster centers. For each feature space, the nearest neighbor is retrieved. The results may not be the same for each descriptor due to noise, pose and the user making slight variations in the way the gesture is performed. We view this as a voting problem with 6 voters where each vote is a recognized gesture. We noticed that the majority of descriptors agree on the correct gesture, so we use the majority rule voting scheme to pick the "winner" candidate gesture. In the case of a tie or if all are different, we classify the gesture as unknown.

4.3 Evaluation

To evaluate the effectiveness of our approach, we designed the following experiment: first we build a training set by asking users (co-authors) to perform the 10 gestures in an arbi-

trary order and repeat the process 3 times. We collected a total of 30 data points per user in each of the feature spaces. In the evaluation sessions, training data was limited to the user testing the system. We did not mix training data among users because of the wide variation in the way different people perform a single gesture. Furthermore, the system is intended to work with hand deformities and restricted joint movements.

After training data was collected, users were asked to perform gestures as prompted by the application for 6 seconds with 3 second breaks between gestures. During the break, the application prompts the next gesture that is to be evaluated. We recorded the following information for every frame: ground truth gesture, majority rule recognized gesture, recognized gesture for each descriptor, time in milliseconds to process raw kinect data (transformations, segmentation, rotation based on principal component) and finally angles α and β found as described in Section 4.2.2. We collected angles to determine whether there is a correlation between hand orientation and recognition performance.

4.3.1 Training Data

The training process consists of computing a set of descriptors Ψ and ψ for each gesture, for a total of 6 data points. We asked users[1] to participate in the training process by performing the 10 gestures in a specific order. They repeat the 10 gestures 3 times with pauses in between. The effect of an arbitrary ordering is more variation compared to repeating the same gesture several times before moving on the next. To demonstrate this variation, we reduced dimensionality using PCA for each descriptor to visualize the clusters they form, see Figure 4.3. Notice how similar gestures form clusters that are closer.

As an example, gestures 5 and 6 ("Palm Front" and "Attention") are similar, as seen in the graphs. Gestures 2 and 5 ("Point Straight" and "Palm Front") are significantly different, placing them farther apart in the feature space.

To further illustrate the effectiveness of the descriptor, we compute distances between

[1]The users for this study were the authors. We are preparing an IRB application for a wider study.

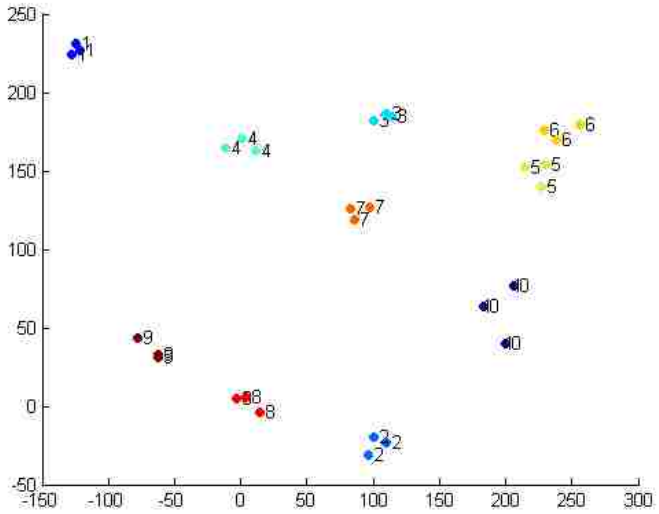


Figure 4.3: Each of the above points is a gesture in the set Ψ_{6^3}. We used PCA to reduce dimensionality and show the well separated clusters in the feature space.

two random sets of gesture examples and graph the results in Figure 4.4. As we expected, the main diagonal has low values, which means that our descriptor is discriminative in ideal conditions. In practice, we have to take into consideration the limitations of the Kinect sensors. As an example, the minimum depth sensing distance is 3 feet. If a user gets closer than the minimum distance, the result is points being dropped from the cloud, which negatively affects the recognition system. Furthermore, because the depth sensor is based on a pinhole camera with inexpensive optics, there is more distortion near the edges of the depth map. In our setup, unless users are restricted to a relatively small space, the detected hand may be near the edge of one or both of the depth maps, which also negatively affect the recognition performance. In the experiments below, we evaluate the performance of our suggested approach by performing 3 experiments. In the first experiment the user does not move the hand, in the second one the user moves her hand arbitrarily and in the third experiment, the user moves her hand in a plus like pattern.

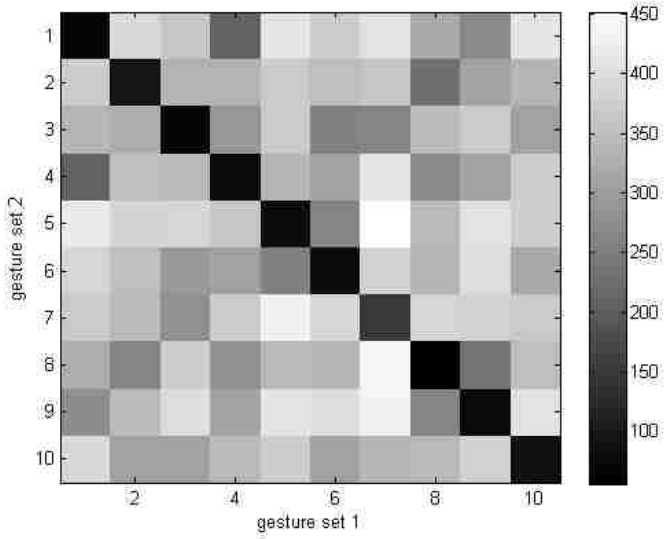


Figure 4.4: Average distances between pairs of gestures using Ψ_{6^3} descriptors.

4.3.2 Experiments

Experiment with user not moving her hand In this experiment, the user performed a set of 10 gestures without moving her hand. Figure 4.5 shows the comparison between the descriptors. The thick red line represents the majority rule gesture accuracy. There were no incorrect classifications in this experiment over a total of 1477 frames collected. We recorded the average α and β for this experiment and graphed α in Figure 4.6. Higher normalization angles resulted from the bias that some gestures cause; e.g., gestures 5 and 6.

We also recorded the average processing time for each frame. The processing times are inclusive of raw data acquisition and transformation, segmentation and recognition. The average processing time per frame was found to be around 35ms on a dual core Pentium D processor running at 2.8 Ghz, without any processor specific optimizations. We wrote the prototype in C#.

Experiment with user making a circle pattern while gesturing This experiment is similar to the first, except the user was asked to move her hand arbitrarily. The total number

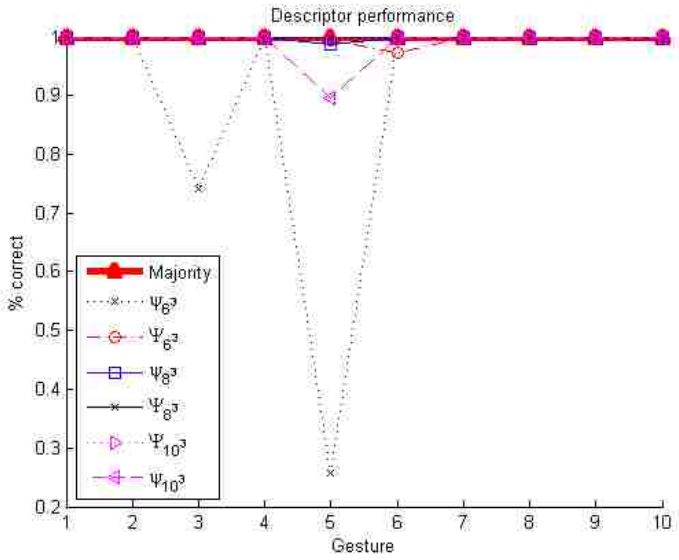


Figure 4.5: The performance of the six descriptors is graphed. Most descriptors performed well, while the majority vote achieved 100% recognition rate.

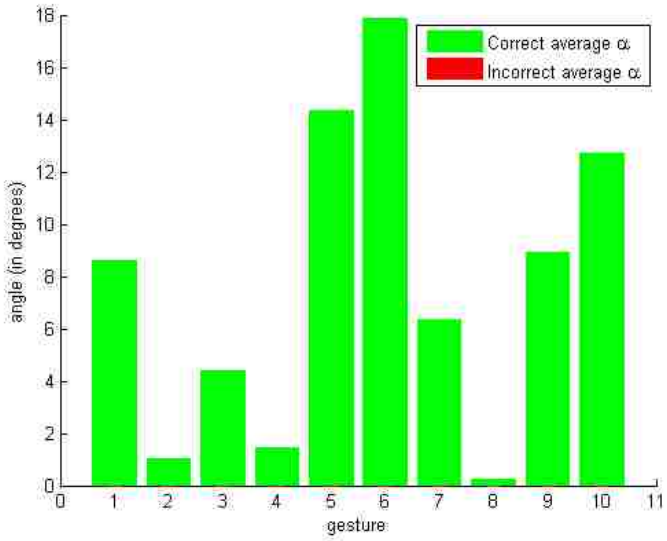


Figure 4.6: The average orientation angle α is graphed for both correct and incorrect classifications by majority rule.

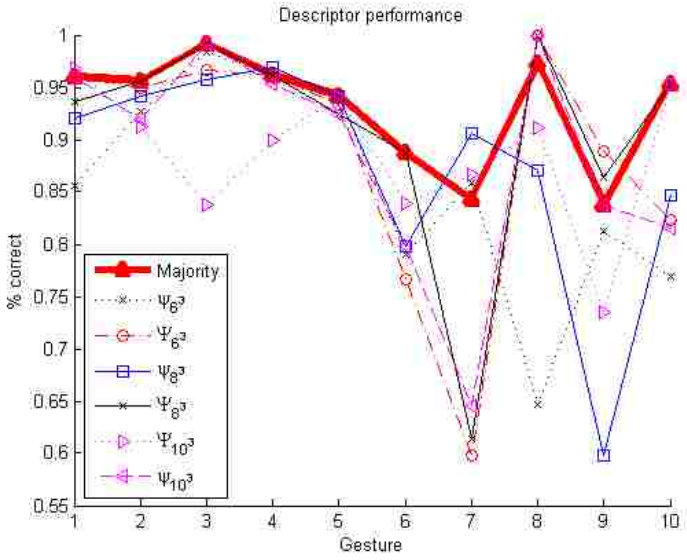


Figure 4.7: We show the performance of the six descriptors when the user moves their hand in a circular pattern. The recognition rate is lower due to orientation normalization.

of frames processed for all gestures was 1274, out of which 1187 were classified correctly (93.17%). Figure 4.7 shows the recognition accuracy of individual classifiers and the majority rule vote. There is a weak correlation between higher normalization angles and recognition accuracy as seen in Figure 4.8. We present the conditional probability distribution of misclassification given angle α in Figure 4.9.

Experiment with a plus like motion In this experiment, the user was asked to perform the set of gestures as in the first two experiments, but he was asked to move his hand in a plus like pattern. We show a confusion matrix in Table 4.1. In Figure 4.10, the performance of individual descriptors and the majority rule is graphed. A total of 1654 frames have been processed, out of which 1602 have been correctly classified (96.86%).

Comparison to other methods We evaluated the effectiveness of our system under various rotation and motion conditions. Similar work has been done by Suryanarayan et al. [97], however, there are differences that prevent a full comparison. First, they use a single time of flight (TOF) camera to augment the 2D images in their gesture vocabulary with

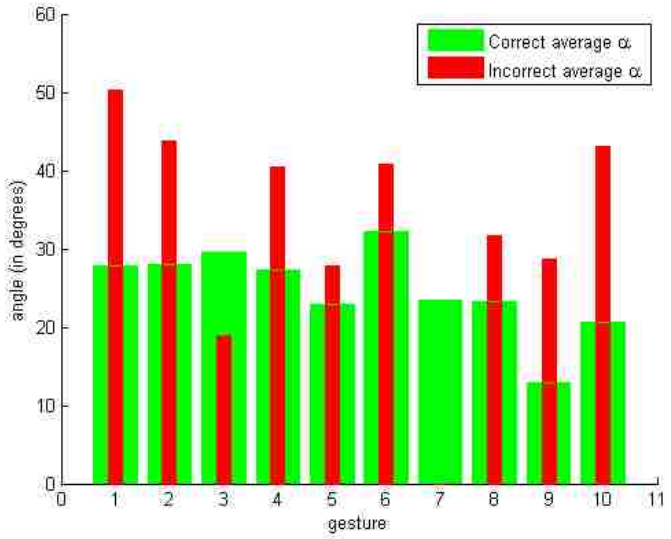


Figure 4.8: The detected average orientation angle α is graphed for correct and incorrect classifications. Note that there is little correlation between normalization angles and recognition accuracy. Graph for angle β is similar.

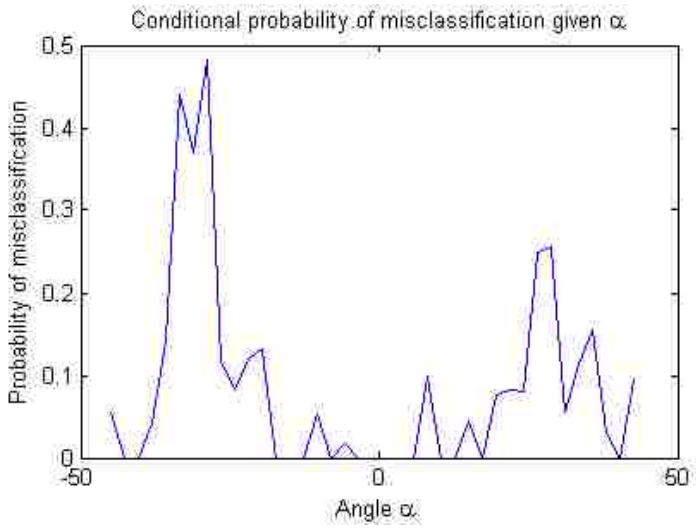


Figure 4.9: We show the conditional probability of misclassification given the angle α. We note that rotation about the Y-axis of the world coordinate system produces the most classification errors.

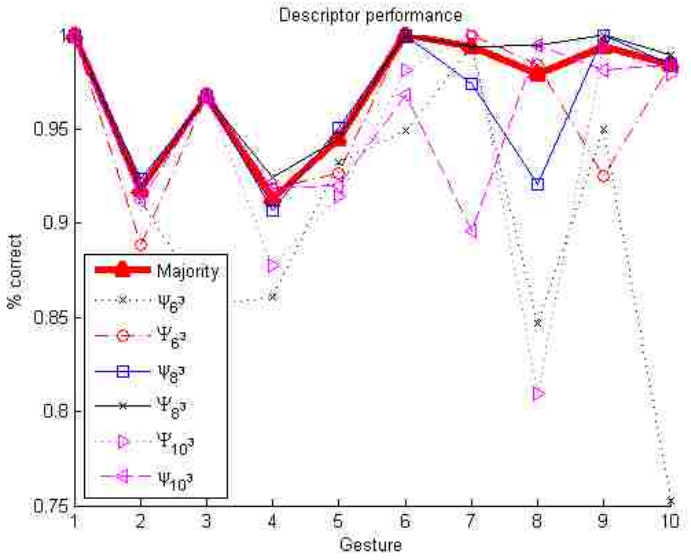


Figure 4.10: Descriptor performance for the third experiment.

Table 4.1: Confusion table. Each row represents a gesture that was evaluated.

Gesture	1	2	3	4	5	6	7	8	9	10	Unknown
1	0.918	0	0	0	0	0	0.005	0.005	0	0.023	0.046
2	0	0.967	0	0	0	0	0	0	0	0	0.032
3	0	0.005	0.912	0.058	0.005	0	0	0	0	0	0.017
4	0	0	0	0.944	0	0	0	0.012	0	0.024	0.018
5	0	0	0	0	1.000	0	0	0	0	0	0
6	0	0.006	0	0	0	0.993	0	0	0	0	0
7	0	0	0.015	0	0	0	0.978	0	0	0	0.005
8	0	0	0.006	0	0	0	0	0.993	0	0	0
9	0	0	0	0	0	0.005	0	0	0.984	0	0.010
10	0	0	0	0	0	0	0	0	0	1.000	0

depth data. They evaluate their system using a vocabulary of 6 gestures, however, it is unclear the motion and rotation constraints they imposed on their evaluation data. They average slightly under 90% accuracy across all 6 gestures. Our system accuracy is 100% when no motion/rotation is performed. We analyzed rotation effects and found them to decrease performance. Our system outperforms the one proposed by Suryanarayan et al. [97] under mild rotation conditions. Motion will decrease performance when the user moves his/her hand fast enough to pass the limitation of the devices' frame capture rate of 30 per second.

4.4 Conclusion

In this chapter we presented a novel algorithm for the recognition of static hand poses using two Kinect sensors. We describe how to merge two depth streams from sensors placed at an angle which captures 3D points belonging to a user's hand. We used a trivial segmentation method for the extraction of hand points and a used PCA to achieve rotation invariance with respect to the X and Y axis of the world coordinate system. The algorithm relies heavily on a clean segmentation. We use multiple descriptors computed from the point cloud and apply majority rule voting scheme to improve the recognition process. The algorithm is evaluated and we present detailed results that can be used for future work in gesture recognition using Kinect sensors. We found that hand orientation has a slight negative impact on the recognition performance of our algorithm, hence a shortcoming. This is mainly due to the segmentation process and distortion near the edges of the depth maps. In future work, we will improve the segmentation process to allow for a less constrained interaction environment.

Chapter 5 A CRF-based Approach to Fitting Generalized Hand Skeleton Models

This chapter has appeared in the written proceedings of the IEEE Winter Conference on Applications of Computer Vision (WACV) 2014 [70]. We present a new point distribution model capable of modeling joint subluxation (shifting) in rheumatoid arthritis (RA) patients and an approach to fitting this model to posteroanterior view hand radiographs. We formulate this shape fitting problem as inference in a conditional random field. This model combines potential functions that focus on specific anatomical structures and a learned shape prior. We evaluate our approach on two datasets: one containing relatively healthy hands and one containing hands of rheumatoid arthritis patients. We provide an empirical analysis of the relative value of different potential functions. We also show how to use the fitted hand skeleton to initialize a process for automatically estimating bone contours, which is a challenging, but important, problem in RA disease progression assessment.

5.1 Introduction

Imaging of the hand is routinely done to diagnose and asses the severity of diseases that alter the normal appearance of the musculoskeletal system. One such disease is rheumatoid arthritis (RA), a chronic systemic inflammatory autoimmune disease that primarily affects joints. The symptoms are pain, swelling and the loss of the joint function due to inflammatory processes. The underlying cause of RA is multifactorial [64, 25] including genetic susceptibilities, nutrition, lack of exercise and environmental factors. Joint inflammation caused by RA leads to over-vascularization, proliferation and synovial scar formation. The synovial proliferation is most marked at the margins of the joints, where the ticht space leads to bone erosions [13]. The inflammatory processes do not spare the ligaments, tendons and muscles, which leads to weakness, laxity and deformity.

We propose a novel method to automatically fit a skeleton model to a hand radiograph.

Our approach builds on a previous model by Fernändez et al. [66], who proposed a point distribution model with landmarks located at joint centers. To support the types of deformation common in RA, we relax this model by adding additional landmark points. Instead of a single point per joint, we have one located on each of the cortical articular surfaces of adjacent bones in the joint. This modification allows us to model subluxation (dislocation) of bones and supports our long-term goal of automatically measuring inter-joint spacing.

We provide a probabilistic formulation of our approach as a Conditional Random Field (CRF) and show to perform learning and inference with the model. We use a collection of potential functions, each tuned to a particular anatomical feature, such as upper and lower joint surfaces, or bone orientation. Each of these features makes unique contributions to our fitting process. We evaluate these features, and our CRF model, on real data from hands with and without deformation due to RA. Based on an analysis of the relative value of different potential functions, we find that the term that estimates the orientation of the joint makes significant contributions to rough alignment but other terms, such as the upper and lower joint potentials, make significant contributions by enabling more precise positioning of landmark points.

The main contributions of this work are: 1) introducing a new point-distribution model suitable for deformed hands, 2) a CRF framework for fitting this model to hand radiographs, 3) the definition of a set of potential functions that focus on specific anatomical structures in the hand, and 4) the evaluation of the accuracy of the method and the relative value of various potential functions on two datasets of hand radiographs.

5.2 Related Work

In this section, we describe previous work on vision-based methods for processing and analyzing hand radiographs.

Hand Radiograph Model Fitting Registering a parametric model to a hand radiograph is a key problem in this domain. Fernändez et al. [66] used a landmark-based wire model (which we extend in this paper) to develop a registration algorithm that outperforms thin-plate splines (TPS). They initialize the wire model through a cascade of image processing routines based on bone axes. Van de Giessen et al. [104] developed a method to register CT scans of wrists by enforcing distances between bone surfaces to remain the same after registration. Bellerini et al. [8] use snakes optimized as initially proposed by Kass et al. [50] using genetic algorithms. They encode the parametric snake as polar coordinates centered at the origin which can be placed arbitrarily on the image. Xu et al. [111] introduced gradient vector flows as external forces, which eliminates the need to know a priori whether the snake will shrink or grow. Our work extends this line of research by generalizing the constrained shape model of Fernändez et al. [66]. This modification leads to the need for improved image feature extraction.

Hand Radiograph Pixel Labeling Numerous approaches have been proposed for pixel-level labeling of hand radiographs, we present several recent examples. Yuksel et al. [112] use a combination of feature classification and morphological operations to segment bone tissue from hand radiographs. Chai et al. [18] use the gray level co-occurrence matrix to segment texture and segment bone tissue from soft tissue. These approaches are similar to our feature extraction approach, but our features were developed to directly aid in model fitting, whereas these were developed for other purposes.

Hand Radiograph Analysis Hand radiographs are used frequently in medical diagnosis because the hands are where the pathology is most evident. We introduce several common medical uses for hand radiographs, each of which could benefit from the improved skeletal model fitting method we propose.

Langs et al. [55] presented a combination of active shape models and active contour models to segment bones and detect erosions on RA patient hand radiographs. Their ap-

proach relies on an initialization based on a local linear mapping net. We point out that hand radiographs of late stage RA patients are much more challenging due to severe subluxation, which result in overlapping bones, and joint space narrowing which lead to weak edge information, thus decreasing performance of purely edge-based methods.

In pediatric radiology, skeletal age is an important indicator of a healthy development process. Not only the bone locations and contours are of interest; bone density measurements aid in the diagnosis of skeletal development. One of the first complete descriptions of a system for hand radiograph analysis for skeletal age assessment is presented by Michael et al. [69]. Hue et al. [46] proposed an algorithm to segment hand bones on hand radiographs of children. Their approach relies on an oversegmentation using the watershed algorithm and region of interest extraction and merging algorithms to segment soft tissue and background from noise. Sotoca et al. [95] proposed a semi-automatic approach where a user places the template at or near the center of a bone and the contour is approximated using active shape models (ASM).

Radiologists rely on expertise to assign a bone maturity score relative to age and gender. The most commonly used method to perform this evaluation is the atlas matching method by Greulich and Pyle (GP method) [8]. This is a time consuming process and correct assessment is highly dependent on the radiologist's experience and expertise, thus automated methods have been proposed. Giordano et al.[36] developed a method to predict bone age using a combination of filtering and Gradient Vector Flow Snakes with accuracy of 90%. Bayesian networks have been used by Mahmoodi et al. [62, 61]. Fuzzy systems have been used for skeletal age assessment by Aja-Fernändez et al. [4].

State-of-the-art bone segmentation and joint space width measurement approaches rely on landmark detection algorithms, usually based on a cascade of image processing techniques. This first step of landmark detection leads to most failures in existing algorithms. Our work fills that gap by accurately computing key anatomical points. roughly centered, currently done manually. Recent work by Davis et al. [24] provides encouraging results on

automating this process.

5.3 Problem Definition

Given a roughly centered hand radiograph, our goal is to estimate landmarks point locations on the edge of cortical articular surfaces along the main axis of long bones. In this section, we formally define our shape model and identify key challenges in solving this problem.

5.3.1 Shape Model

We represent a shape, s, by a set of n landmark point locations:

$$s = (x_1, x_2, x_3, ..., x_n, y_1, y_2, y_3, ..., y_n)^T.$$

The choice of landmarks depends on the object of interest and the application, but for hand radiographs they are usually chosen as joint centers and fingertips. A recent example is the work of Fernändez et al. [66] where a shape model is used as an initialization step to an image registration algorithm. We chose to generalize their representation by having two landmarks per joint, one on each side of the joint on the cortical articular surface of the bone, collinear with the bone's main axis, (i.e., the tips of long bones). Figure 5.1 shows a visualization of this model. This relaxation allows us to model the subluxation deformities that are common in moderate to late stage rheumatoid arthritis patients.

5.3.2 Key Challenges

Automated methods for radiograph analysis rely on consistent alignment and appearance, which rarely happens in practice. In this work, we focus on solving the alignment problem by fitting an initial hand model to the radiograph. We describe several important challenges in solving this fitting problem.

Despite attempts to control hand position using clinical protocols, hand radiographs of healthy patients show significant variations in pose. In addition to RA deformities, such

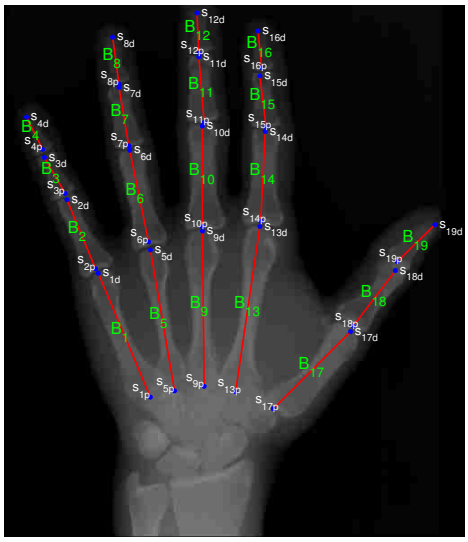


Figure 5.1: Our proposed shape model: each segment corresponds to a bone ($B_{1...\ 19}$). Individual points are indexed as proximal $s_{\{1...19\}p}$ and distal $s_{\{1...19\}d}$.

as joint fusion, sclerosis and subluxation, other diseases including osteoarthritis may be present, further increasing variability in pose. For example, RA damage and pain can prevent patients from flattening their hands on the imaging surface. A solution to the hand model fitting problem must be able to cope with significant changes in pose and joint deformities.

The appearance of bones also varies significantly from patient to patient. This is especially true in RA patients because the disease affects the density and shape of individual bones. Many previous approaches to the hand model fitting problem focus on specific anatomical features for alignment, but this leads to brittle solutions. Therefore, an approach that combines image information from multiple anatomical features is needed. Our work addresses both of these concerns in a consistent, and adaptable probabilistic formulation.

5.4 Approach

We propose a CRF-based model that combines a shape prior with appearance terms that identify various anatomical structures. The shape prior is based on the distance from the subspace spanned by a Point Distribution Model (PDM) and the appearance is defined as

a collection of likelihood terms that depend on local feature detectors. We use a local optimization strategy to jointly maximize feature responses at landmark locations by minimizing an energy function. Minimizing energy in this context is equivalent to maximizing the posterior distribution over the correct location of landmarks given an image.

Our proposed CRF has the following form:

$$P(s|I,\theta) = \frac{1}{Z} exp\{\sum_i \{\sum_j \Psi_j(s_{i_p}, s_{i_d}, \theta)\} + \\ + \sum_k \phi_k(s_i, \theta) + \zeta(s, \theta)\} \tag{5.1}$$

where i is an index over model segments, j and k index our pairwise and unary appearance terms, Z is the partition function, Ψ and ϕ are pairwise and unary appearance terms, ζ is a shape model prior and θ is a weight vector we use to balance the various terms.

5.4.1 Potential Functions

The data terms in our model are based on a collection of discriminative features that we combine into a set of potential functions.

Discriminative Features The success of the shape fitting process is heavily dependent on a set of features that are highly discriminative. Recently, Cootes et al. [21] showed how regression voting using random forests in the constrained local model (CLM) framework outperforms existing methods on shape fitting. Our approach extends Constrained Local Models (CLM) [22], by formulating the shape fitting problem in a general CRF framework.

We classify each pixel in an image independently using a randomized decision forest (RDF) classifier and use dense SIFT features as input. We train 4 RDF classifiers, one for joint centers (the midpoint between adjacent bones connected by joints), two for cortical articular surface points, one for proximal and one for distal, and a bone tissue classifier.

For a new image, we compute DSIFT descriptors and run them through our RDF classifiers. The output is a likelihood that represents class membership of each pixel. We note

that this process implies independence in pixel memberships (e.g., we could potentially have a pixel be labeled as both bone and cortical surface). Let the classifier response for joints be $f_j(s_i)$ where s_i is a landmark point with a corresponding image location. Similarly, we define $f_{pc}(s_i)$ and $f_{dc}(s_i)$ for proximal and distal cortical articular surface pixels. Finally, let $f_b(s_i)$ be the classifier response for bone tissue. Examples of RDF classifiers outputs can be seen in Figure 5.3.

We then apply a thresholding operation on f_j and f_b to compute binary regions of high probability bf_j and bf_b. Using the thresholded responses, we apply distance transformations, and combine them with values inside the regions to compute $df_j(s_i)$ and $df_b(s_i)$. The distance transformation returns 1 on the region borders and 0 inside. The locations inside the regions are filled with $1 - f_j$ and $1 - f_b$ respectively. This approach increases alignment precision by providing extra information about the most probable location of anatomical interest points. Examples can be seen in Figure 5.4.

In the next section we show how we convert these low-level image features into potential functions in our CRF model.

Pairwise potentials The first pairwise potential encodes the compatibility between a segment in our shape model and evidence of bone tissue:

$$\Psi_1(s_{i_p}, s_{i_d}) = \theta_1 \frac{1}{n} \sum_{n=1}^{t} df_b(p_{x_n}, p_{y_n}) \tag{5.2}$$

where the summation is over points p sampled along a segment. This function takes as input the distance transformation df_b and is low when a segment is placed over a bone.

The bone evidence from the image can be further exploited by considering segment orientation extracted from the thresholded bone tissue classifier connected components. We define the following potential:

$$\Psi_2(s_{i_p}, s_{i_d}) = \theta_2 \, (tan^{-1}(s_{i_p} - s_{i_d}) - f_o(s_i))^2. \tag{5.3}$$

In the above equation, f_o is a function that returns an angle at an image location computed via a weighted averaging of angles of the connected components with respect to the hori-

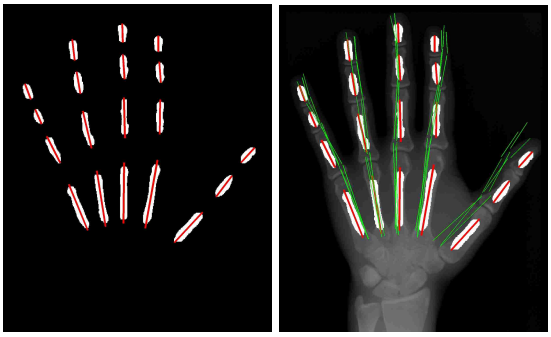


Figure 5.2: Left: in red, segments span the major axis of connected components in a binary image (bf_b) used for initialization. Right: in green, the top 3 models from the training set with the lowest ICP registration error.

zontal image axis. Intuitively, if a segment s_i is placed perpendicular to the major axis of a connected component, the potential will be at its maximum. In Figure 5.4 we show f_o for an image.

We now define a pairwise potential that encodes a prior over adjacent cortical articular surfaces in a joint:

$$\Psi_3(s_{i_d}, s_{i+1_p}) = \theta_3(-log\,\mathcal{N}(0, \Sigma)) \tag{5.4}$$

where $\mathcal{N}$ is a Gaussian with full rank covariance Σ computed from the training set. This term constrains joint spaces to be at reasonable distances in order to avoid local minima during optimization.

Unary potentials We define three terms that encourage the landmark points to be near appropriate anatomical features of the joint. The motivation for our first unary potential is that all points on the model should be in areas of high probability indicated by our joint feature, f_j. To further penalize points from being far from a joint, and to improve optimization performance, we use df_j, which is an augmented version of f_j. This unary potential is defined as follows:

$$\phi_3(s_i) = \theta_6\, df_j(s_i). \tag{5.5}$$

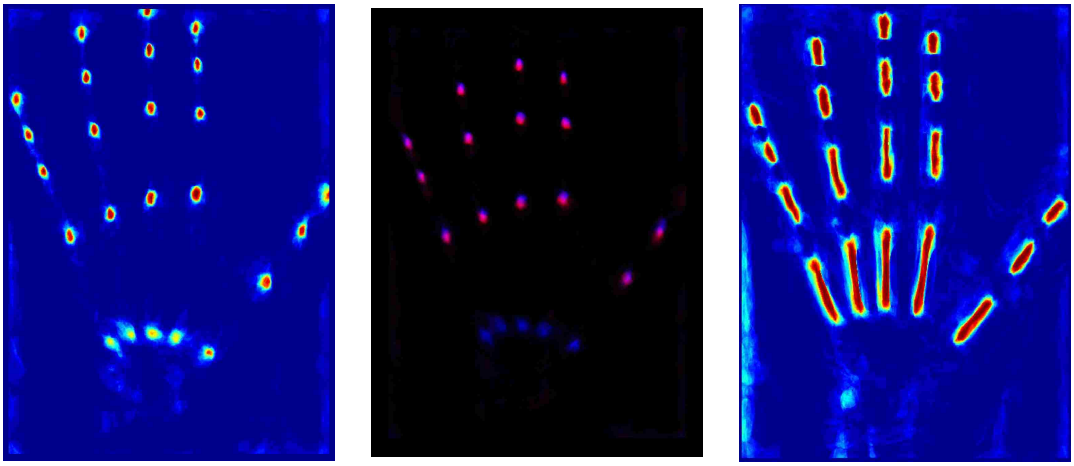


Figure 5.3: Left: joint center pixel probabilities. Middle: color coded probabilities for distal (red channel) and proximal (blue channel) cortical surface pixels. Right: bone pixel probabilities.

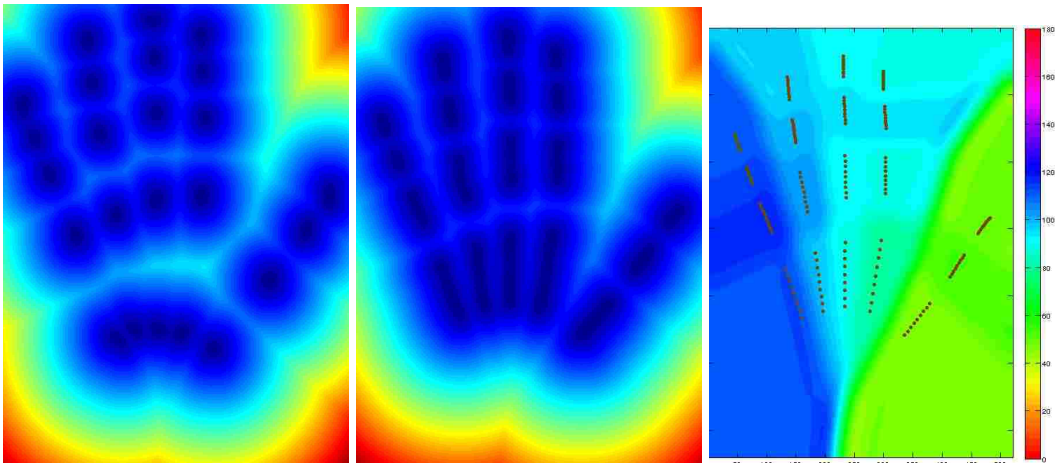


Figure 5.4: Left and middle: Distance transformations $df_j(s_i)$ and $df_b(s_i)$ of thresholded classifier responses. Right: Orientation term f_o.

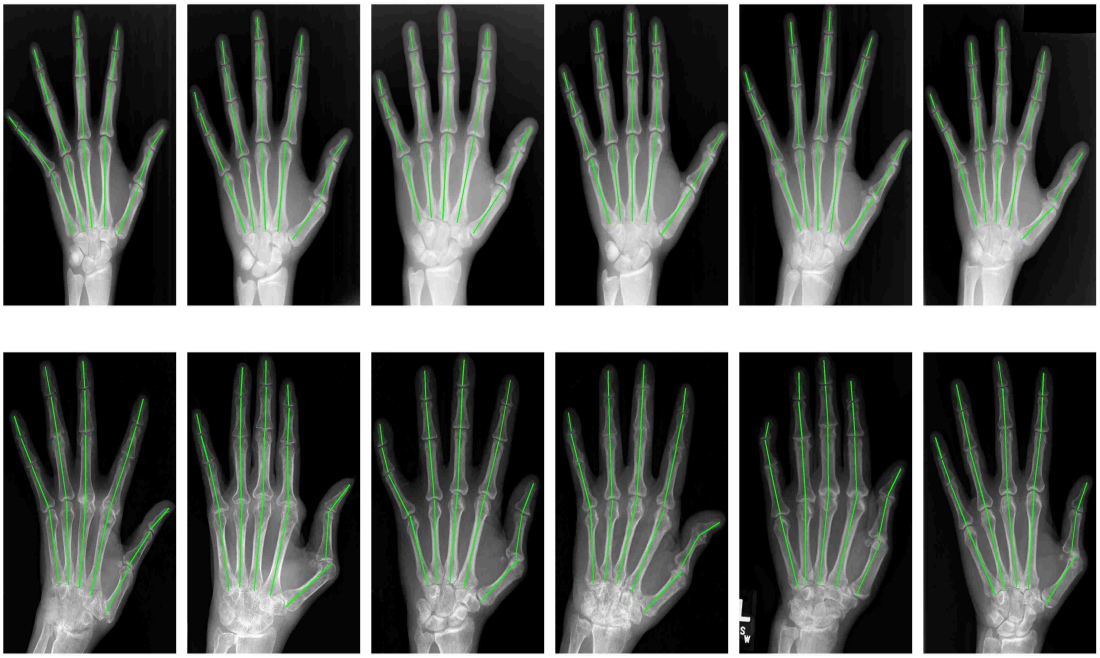


Figure 5.5: Results on a representative subset of test images. Top row: healthy radiographs from the Hand Atlas Database. Bottom row: rheumatoid arthritis set.

The second two unary potential functions encourage landmark points to align to the joint contours. These potentials are defined as follows:

$$\phi_1(s_{i_p}) = \theta_4 \; f_{pc}(s_{i_p}). \tag{5.6}$$

and

$$\phi_2(s_{i_d}) = \theta_5 \; f_{dc}(s_{i_d}). \tag{5.7}$$

For this domain using classifiers, f_{dc} and f_{pc}, instead of a generic edge detector is critical because it allows the model to distinguish between the true bone contours used for diagnosis and analysis and apparent edges caused by the radiographic projection of other bone structures.

We find that in practice these terms complement each other. The first is very important for rough initial alignments, while the second two are critical for the precise alignments. See the evaluation section for details.

5.4.2 Shape Model Prior

The shape prior term $\zeta(s, \theta)$ is used to penalize unlikely shapes. Using Probabilistic Principal Component Analysis (PPCA), we seek to relate shape s to a k-dimensional vector x that is normally distributed with zero mean and covariance $I(k)$:

$$s^T = Wx^T + \bar{s} + \epsilon \tag{5.8}$$

where W is the matrix of principal components and $\bar{s}$ is the average shape. ϵ is the model noise component, assumed to be normally distributed $\epsilon \sim \mathcal{N}(0, \sigma^2 I)$.

Under this model, s is normally distributed:

$$P(s) = \mathcal{N}(\bar{s}, WW^T + \sigma^2 I(k)) \tag{5.9}$$

so our shape prior is a weighted negative log likelihood of $P(s)$:

$$\zeta(s, \theta) = \theta_7(-log\ P(s)) \tag{5.10}$$

W and σ^2 are estimated using an Expectation-Maximization algorithm from a training set of hand shapes extracted from hand radiographs.

5.4.3 Shape Inference

We now describe our strategy for estimating the optimal shape model for a given hand radiograph. We first compute a set of initial models, then use local optimization for each and select the best.

We use thresholded bone tissue classifier, bf_b, for initialization by computing the connected components statistics in the binary image. We model each component as an ellipse and compute its centroid, major axis, orientation and length. The line segments spanning the major axis of the connected components (see Figure 5.2) form the basis of our initialization scheme. We use a variation of the Iterative Closest Point algorithm to register samples from our training set to the segments extracted from the binary bone tissue classifier. The

ICP registration error is then used to select the 3 best shapes that are used as starting points for our local optimization.

Our local shape objective function is the posterior probability of a shape, s, given image data, or equivalently the log likelihood of our CRF model (5.1), which is defined as follows:

$$E(\hat{s}, \theta, I) = \arg\min_{s} \sum_{i} \{\sum_{j} \Psi_j(s_{i_p}, s_{i_d})\} + \sum_{k} \phi_k(s_i) + \zeta(s, \theta). \tag{5.11}$$

To minimize (5.11) we use coordinate descent with step sizes chosen by an independent local search in each dimension.

5.4.4 Estimating CRF Weights

We use a supervised learning approach to estimate the model parameters θ. The partition function Z is NP-hard to compute [53]. We overcome the difficulty of computing Z in learning the model parameters by using Pseudo-Likelihood learning, where a uniform prior over model parameters is assumed by setting $\tau = \infty$ in $P(\theta|\tau) = \mathcal{N}(\theta, 0, \tau^2\mathbf{I})$ where I is the identity matrix. We find the optimal parameters θ by minimizing the difference between our estimated shapes and ground truth shapes over a set of training images:

$$\hat{\theta}^{ML} = \arg\min_{\theta} \sum_{i} \left\| \arg\min_{s} E(s, \theta, I_i) - s_i^{GT} \right\|_2^2. \tag{5.12}$$

The above minimization is non-convex since we allow s to vary during optimization. We use the simplex method with random restarts to compute model parameters, θ. In the following section we evaluate the model parameters and show results from inference.

5.5 Evaluation

Datasets We evaluate our method on two datasets with posterior-anterior view hand radiographs: the Digital Hand Atlas Database[1] and a set of 43 radiographs of RA patients from the University of Kentucky Department of Radiology. The second dataset is of a

[1] http://www.ipilab.org/BAAweb/

hands in a range of disease stages, from minimal to extreme deformation. For evaluation purposes, we manually determined landmark locations for each image in both datasets. Since radiograph contrast varies due to calibration parameters and noise, we truncate the upper 20% of the intensity histogram.

Quantitative Evaluation We used a set of 20 images from both datasets to train our discriminative classifiers and estimate our remaining model parameters. The remaining images were used for testing and validation.

We evaluate the model by computing the sum of absolute differences between ground truth shapes and results from inference. The model errors for both datasets can be seen in Figure 5.7 and 5.6. For the Hand Atlas set, the average per point error was 2.72 pixels, while for the RA dataset it was 2.85 pixels. A comparison to the state of the art is difficult due to our model landmark selection and RA deformity severity. We divide the test set into early (16 images), moderate (11 images) late stage radiographs (11 images), the average per point errors (measured in distance from ground truth, in pixels) are as follows: 2.30, 2.24, and 4.56.

To help understand the failure modes of our approach, we further investigate two radiographs with poor shape estimates. These correspond to images 23 and 27 in Figure 5.7. We find that by inspecting the optimal fit for both images, shown in Figure 5.8, that they are both from patients with late-stage RA and have severe deformities and subluxation. In such cases, assistance from a radiologist will be required.

Term Contributions To provide more insight into the model, we estimate the amount each potential function contributes to reducing errors in the RA dataset. We split the dataset into two groups, a training set of size 20 and a testing set of size 12. We use the training dataset to estimate the optimal potential function weights, $\hat{\theta}$, by minimizing (5.12), as described above, for the full model. Then, for each of the seven potential function, we solve for the optimal set of weights for the model without that potential function, leaving

Table 5.1: Percent increase of error for each term when omitted from the model.

Term	Ψ_1	Ψ_2	Ψ_3	ϕ_1	ϕ_2	ϕ_3	ζ
% Error increase	39.14	1.42	1.87	2.19	0.09	10.56	0.27

one out. This results in a set of eight different models. For each model, we infer the shape in each image in the testing dataset and sum the absolute pixel error with respect to the ground-truth shape to obtain an error measure. Table 5.1 shows the ratio of the error of a model without the potential function to the error for the full model. Intuitively, an important term will result in a model with significantly higher error if it is removed. We find that there are two dominant terms: Ψ_1 and ϕ_3 that correspond to bone and joint evidence from the feature set. These terms are clearly the most important in gross alignment, however the other terms each make a contribution to reducing errors in the full model.

Example Applications: Initialization for Estimating Bone Contours The weakness of most state-of-the-art approaches for identifying bone contours is the initialization step. As an example application, we propose to fit an active shape model (ASM) to the cortical articular surfaces for each finger joint. This is challenging because ASM models must be initialized very close to the optimal location or they will fall into non-optimal local minima. We use our proposed approach to estimate a model skeleton and use the landmark points and bone segments to align the initial ASM model for each joint. We optimize the ASM shape parameters, using an off-the-shelf software library, and obtain the results seen in Figure 5.9. This demonstrates that our estimated skeleton models are sufficiently accurate to provide initial conditions for ASM models of joint contours. In combination, such an approach could be used to automatically estimate the joint space width, an important metric for RA progression.

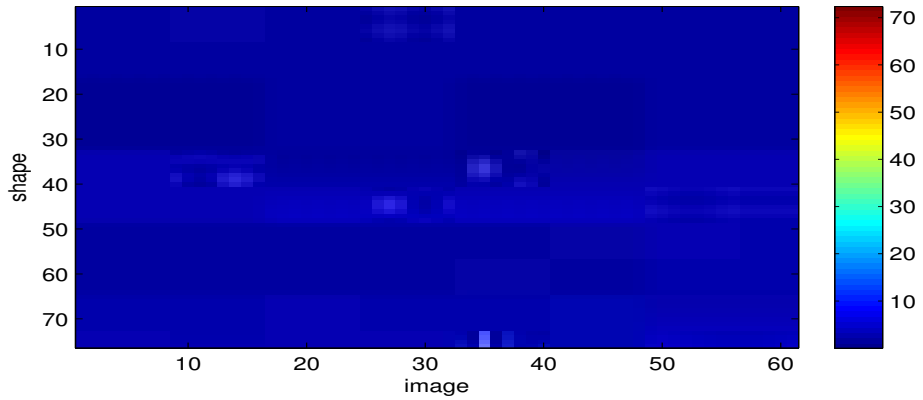


Figure 5.6: Hand Atlas Dataset model errors computed as sum of absolute differences from ground truth.

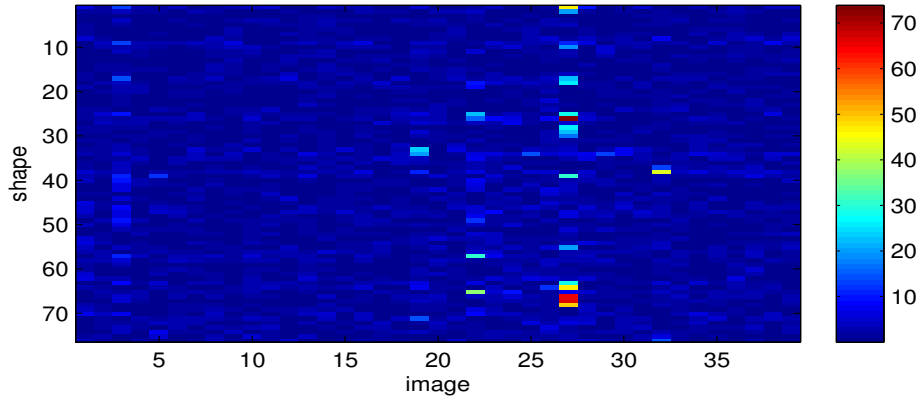


Figure 5.7: Rheumatoid Arthritis Dataset model errors computed as sum of absolute differences from ground truth.

5.6 Conclusion

We introduced a new method for fitting wireframe-hand models to radiographs. A key innovation in our approach is in fitting a relaxed shape model, with four degrees of freedom at each joint, that is capable of representing the dramatic subluxations present in patients with rheumatoid arthritis (RA). Fitting this model effectively is more challenging than standard models, which only have two degrees of freedom at each joint. We show that our method, which combine low-level discriminative features in a conditional random field framework, is capable of fitting this relaxed model on healthy hands as well as those deformed by RA. We provide quantitative results that highlight which features are most important and show

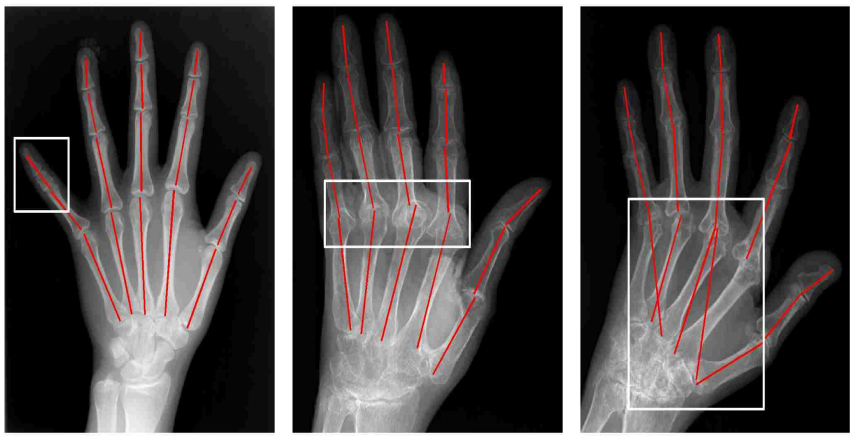


Figure 5.8: Failures that correspond to high errors from Figure 5.6 and 5.7.

an application of our method to fitting bone contours, which is critical in assessing RA damage.

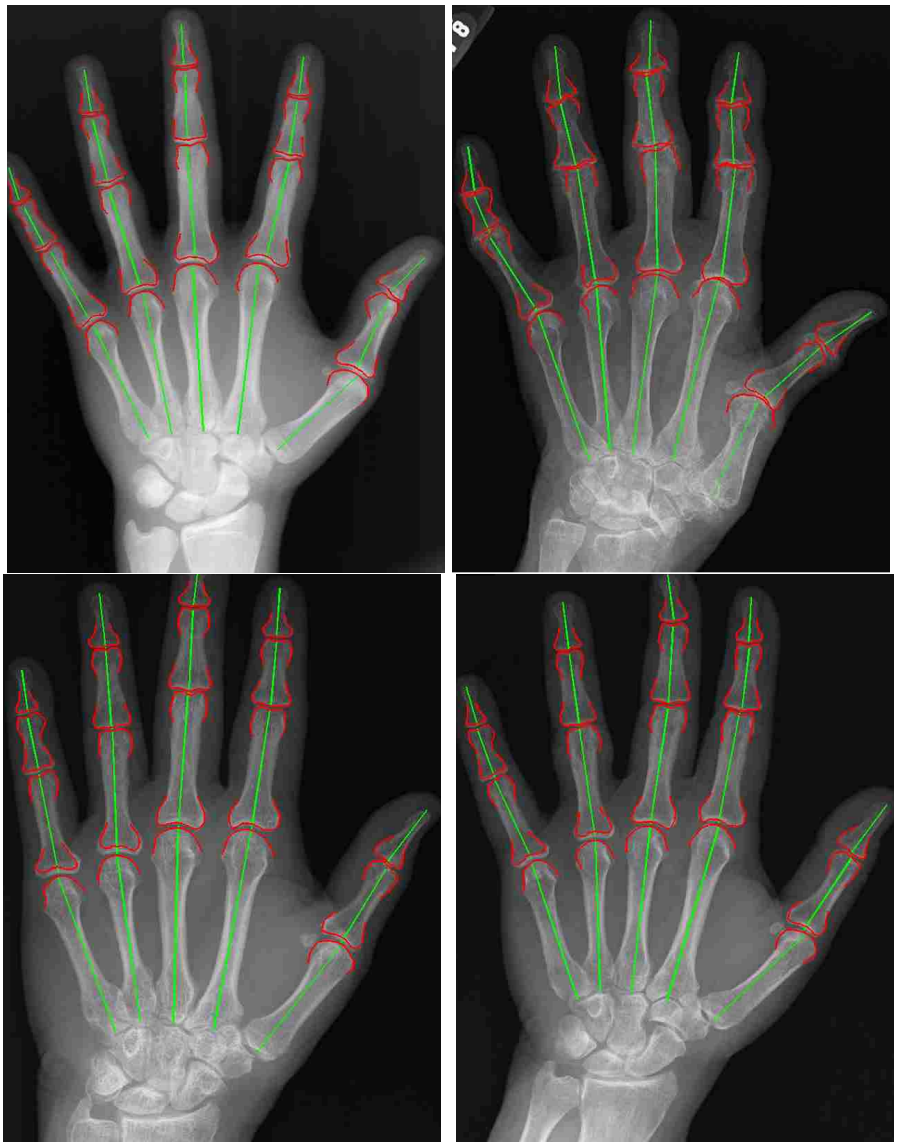


Figure 5.9: Joint contours (red) estimated by initializing an active shape model based on our initial hand skeleton (green).

Chapter 6 Disease Stage Metric Learning

The shape of the hand of an RA patient depends on individual and disease related variability. Examples of individual variability include differences due to age and genetics, while disease related variability includes changes due to degenerative processes of autoimmune responses. In this work, we propose a novel metric learning approach for anatomical shapes. We were inspired by the geometric intuition that distances between all healthy samples should be zero, while distances between diseased samples from the same patient should be proportional to the stage that was assigned by a medical expert. This idea is motivated by datasets of medical imagery in which samples are given a discrete stage label, but the progression is a continuous measure.

In medical practice, hand radiographs are labeled as: healthy, early, moderate and late, as described in Chapter 5. These labels give us partial information about the actual stage, i.e., only the interval. The true stage is a continuous positive value. This problem formulation falls under semi-supervised learning methods [113].

We propose an algorithm to learn a quadratic Gaussian metric for use on anatomical shape representations in regression and classification tasks. Our method was inspired by the observation that anatomical structure variation is the result of two processes: inter-patient (intrinsic) variability and variability due to a disease process (extrinsic). Given this observation, our intuition is that a good distance metric for this purpose is one under which all samples from healthy individuals should be very close to zero, and the distance between diseased samples should be proportional to the stage of the disease.

6.1 Background: Supervised Metric Learning

Learning a good distance metric over an input space is an important problem in the machine learning literature. Distance learning from available data can improve algorithms that rely

on pairwise similarities between data points (e.g., nearest neighbor and support vector machines). Moreover, one can gain valuable insights into the structure of the data by means of visualization and dimensionality reduction. Perhaps the most widely used distance metric is the squared Euclidian distance, which is defined for two sample points x_i and x_j as:

$$D_{ij}^2 = (x_i - x_j)^T(x_i - x_j). \tag{6.1}$$

We note this relationship because of a related problem in machine learning, namely feature extraction, where one seeks a function $f(x)$ where x is a raw data point and the result is a feature vector or a point in some feature space. We denote any distance function in feature space as $d_{ij}[f(x_i), f(x_j)]$. A common class of feature spaces is a linear projection of x: $f(x) = Wx$. This gives rise to the class of Mahalanobis distances:

$$\begin{aligned} d_{ij}[f(x_i), f(x_j)]^2 = & \quad (f(x_i) - f(x_j))^T(f(x_i) - f(x_j)) && (6.2)\\ = & \quad (Wx_i - Wx_j)^T(Wx_i - Wx_j) && (6.3)\\ = & \quad (x_i - x_j)^T(W^TW)(x_i - x_j) && (6.4)\\ = & \quad (x_i - x_j)^TA(x_i - x_j) && (6.5) \end{aligned}$$

where $A = W^TW$ is a positive semidefinite matrix. Our goal is to learn A, such that d_{ij} is a valid metric, i.e., it satisfies the following conditions:

1. $d_{ij}[f(x_i), f(x_j)] > 0$
2. $d_{ij}[f(x_i), f(x_j)] = 0$ *iff* $f(x_i) = f(x_j)$
3. $d_{ij} = d_{ji}$
4. $d_{ik} \leq d_{ij} + d_{jk}$.

The first two conditions imply positive definiteness on A. Conditions 3 and 4 enforce symmetry and the triangle inequality. Our goal in this chapter is to learn A, in a semi-supervised setting.

6.2 Related Work

Common approaches to solve for A take advantage of prior knowledge, usually in the form of equivalence relations that specify which points should be close and which points should be distant. A standard approach in machine learning is Linear Discriminant Analysis (LDA), first explored and proposed by Fisher et al. [30]. LDA can be thought of as a supervised version of PCA, where the variance in the latent space and the inter-class spread is maximized. This method explicitly models the differences between classes of data. Both PCA and LDA assume the distribution of the data in the latent space is Gaussian.

Xing et al. [110] propose a distance metric learning method that takes advantage of user provided pairwise similar points. Their method is formulated as a semi-definite program (SDP). A closely related approach by Li et al. [58] proposes to learn a smooth mapping to a latent space from pairwise "must-link" and "cannot-link" constraints, where "must-link" pairs are mapped to the same point and "cannot-link" pairs are mapped to be orthogonal.

Globerson et al. [37] introduce maximally collapsing metric learning (MCML), a method to learn a Mahalanobis distance metric based on similarity and dissimilarity constraints. Their construction involves a convex optimization problem where samples from one class are "collapsed" into a single point, while samples in other classes are pushed infinitely far apart. The main difference between our method and MCML is that pairwise constraints are used in their construction while we require classes to map close to hyperspheres (circles in 2D, centered at the origin) of radius proportional to a numeric stage label.

Goldberger et al. [38] propose a distance metric learning method specifically designed to improve the KNN classifier. Their cost function is the leave-one-out performance of the KNN classifier on training data. Although our method can be used in combination with a KNN classifier with a similar goal, our method seeks an embedding with the property that magnitude is directly related to variance due to a disease process.

6.3 Our Semi-Supervised Approach

We translate categorical labels to numerical values and intervals as follows $healthy = 0$, $early \in [1, 2)$, $moderate \in [2, 3)$, $late \in [3, \infty)$. We seek a distance metric that gives rise to an embedding where the stage is interpreted as the distance from the origin. Samples in the same discrete stage should be mapped onto the latent space with magnitudes in the intervals as described above.

Let $X = [x_1, x_2, ..., x_N]$ be a set of examples where $x \in \mathcal{X}$ and $Y = [y_1, y_2, ..., y_N] \in \mathcal{L}$ be the stage labels. We seek a distance measure for points in data space, $\mathcal{X}$. More specifically, we seek a Mahalanobis distance of the form:

$$d(x_i, x_j|A) = d_{ij}^A = \sqrt{(x_i - x_j)^T A (x_i, x_j)}, \tag{6.6}$$

where A is constrained to be a positive semidefinite matrix.

When $A = W^T W$ and W has m columns and n rows, it suffices to constrain the rank of W to be n for A to be positive semidefinite. This formulation for A is useful to consider because one can also think of the Mahalanobis distance as a mapping of $x \in \mathcal{R}^m$ to $\mathcal{R}^n$ where $m < n$ via a linear projection Wx $(A = W^T W)$, thus reducing dimensionality. Intuitively, our method solves for W such that the norm of the projection on the subspace defined by it is related to disease stage.

6.4 Disease Stage Metric Learning

We now formalize our disease stage metric learning (DSML) problem as an optimization problem of the following form:

$$\arg\min_W \frac{1}{N} \sum_{i=1}^{N} (x_i^T (W^T W) x_i - y_i^2)^2 \qquad s.t.\ W^T W \succeq 0. \tag{6.7}$$

This formulation encourages samples to map onto rings centered around the origin. The vector magnitudes in the subspace defined by W become functions of the stage label. We

use an off-the-shelf optimizer to compute W. The cost function in Equation 6.7 is non-convex, since the outer square term $x_i^T(W^TW)x_i - y_i^2$ can be negative for some choices of W. The cost function is invariant to rotation, hence the solution space is infinite.

One limitation of the above cost function is that shapes with similar deformities can be mapped arbitrarily far in latent space. Ideally, we should observe clusters of shapes with similar deformations in the latent space. One approach is to add a regularization term to our cost function to encourage samples that are close in the shape space to be close in the embedded space:

$$\arg\min_{W} \frac{1}{N}\sum_{i=1}^{N}(x_i^T(W^TW)x_i - y_i^2)^2 + \lambda|X^TX - X^T(\frac{W^TW}{Tr(W^TW)})X| \qquad s.t.\ W^TW \succeq 0. \tag{6.8}$$

In the equation above, the intuition behind the regularization term $|X^TX - X^T(\frac{W^TW}{Tr(W^TW)})X|$ is that pairwise squared Euclidian distance differences between the raw shape space and latent space should be minimized. λ is a parameter that controls the relative importance of the regularization term.

6.5 Evaluation

For evaluation we show results on the stage classification problem with low rank constructions of A (when W has 2 or 3 columns for visualization). We implemented the algorithm in MATLAB and used the built-in unconstrained optimization routine *fminunc*, which uses a combination of algorithms to estimate the gradient and search.

6.5.1 Linear Generative Data Model

We now formalize the notion of separating sources of variance in shape models. Consider a linear generative model of shapes of the form:

$$s = \mu + \alpha\mathbf{I} + \gamma\beta\mathbf{E} + \varepsilon \tag{6.9}$$

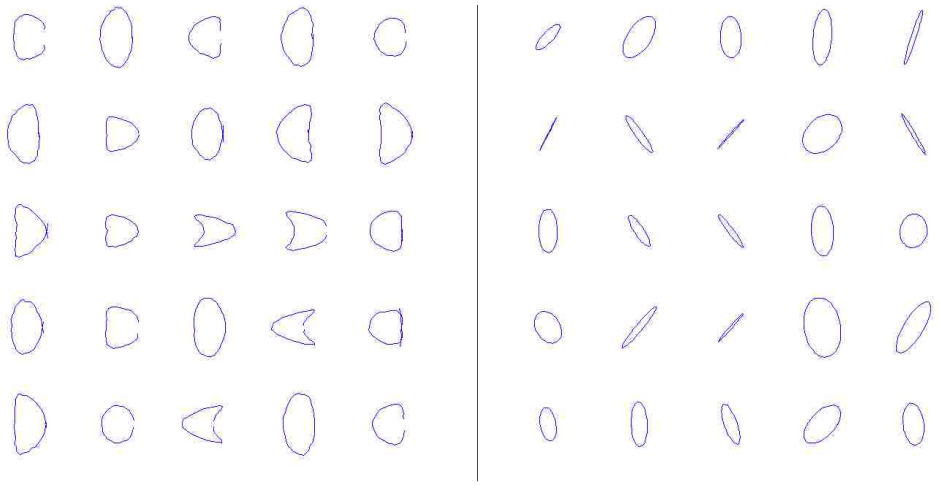


Figure 6.1: When the average shape μ is a circle, we sample exclusively from the intrinsic (left) and extrinsic (right) modes of variation.

where μ is the average shape, the columns of $\mathbf{I}$ define an intrinsic variability subspace (e.g., genetic differences lead to anatomical shape variability), the columns of $\mathbf{E}$ define an extrinsic variability subspace (e.g., deformations due to disease processes), ε is observational noise. We assume Gaussian distributions for $\alpha \sim \mathcal{N}(0, I\sigma_{\mathbf{I}})$ and $\beta \sim \mathcal{N}(0, I\sigma_{\mathbf{E}})$, and $\varepsilon \sim \mathcal{N}(0, I)$. Variable $\gamma \in [0..\infty)$ is a scalar and is used to control stage, $\sigma_{\mathbf{E}}$ is the "stage" noise (e.g., physicians are not always 100% accurate on staging a disease).

By adjusting the variance of α and β we can effectively control the mix of variability due to differences in patients and differences caused by a disease process. In our experiments, we show that our method is effective even when the intrinsic term of the generative model is non-linear (which is the case with real medical data). The strength of our method is most evident when $Var(\alpha) > Var(\beta)$, as we will show in the evaluation.

We now construct a synthetic dataset sampled from our generative model. We define the average shape to be a unit circle and construct intrinsic and extrinsic deformation components. Let $\mu \in \mathcal{R}^D$ be a column vector $\mu = \begin{bmatrix} x^T \\ \\ y^T \end{bmatrix}$. where $x^T = cos(\theta)$ and $y^T = sin(\theta)$ and θ is a set of values sampled at equal intervals from $[0..2\pi]$. We define 2 intrinsic and 3 extrinsic components $\mathbf{I} = [\mathbf{I}_1\ \mathbf{I}_2]$ and $\mathbf{E} = [\mathbf{E}_1\ \mathbf{E}_2\ \mathbf{E}_3]$. To illustrate this model, we sam-

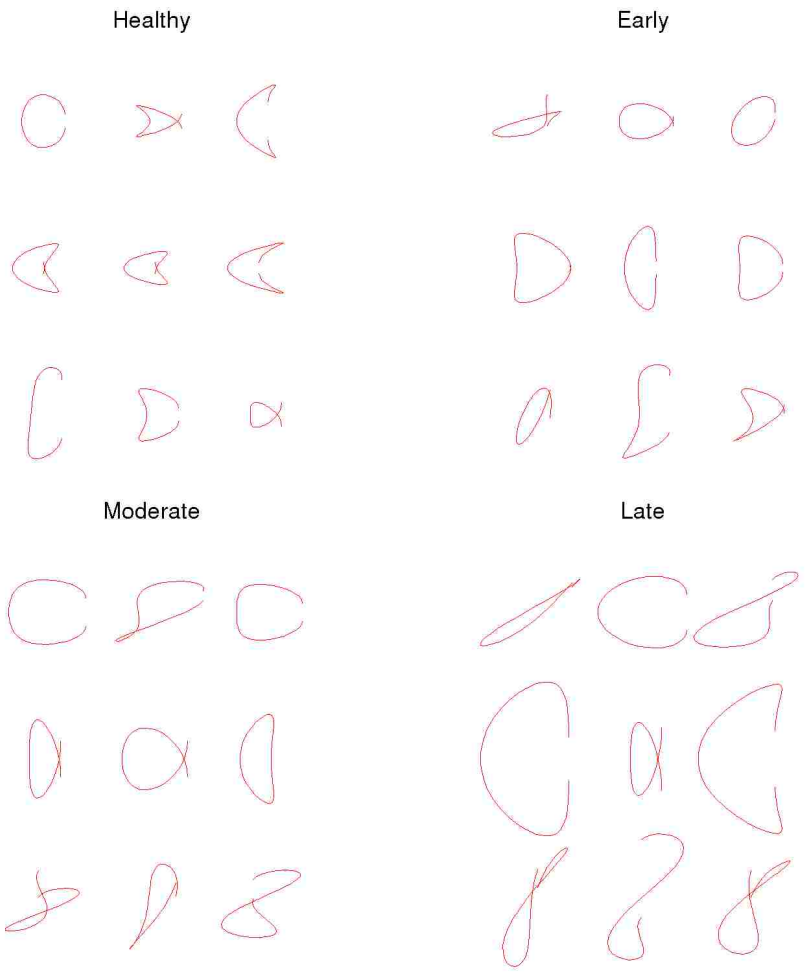


Figure 6.2: We show random samples from our generative model while controlling for stage. The intrinsic and extrinsic component variances are equal at 0.5.

ple separately from the intrinsic and extrinsic components of this model in Figure 6.1. In Figure 6.2 we show the effect of the disease stage parameter γ.

6.5.2 Embeddings using Synthetic Data

We first compare the results of the embeddings produced by our method (DSML) and PCA for several choices of σ_I, σ_E with zero noise. In Figure 6.3 we show the resulting embeddings when $\sigma_I > \sigma_E$ and $\sigma_E > \sigma_I$. The strength of our method is most obvious when the intrinsic variance is higher since it can still isolate the extrinsics.

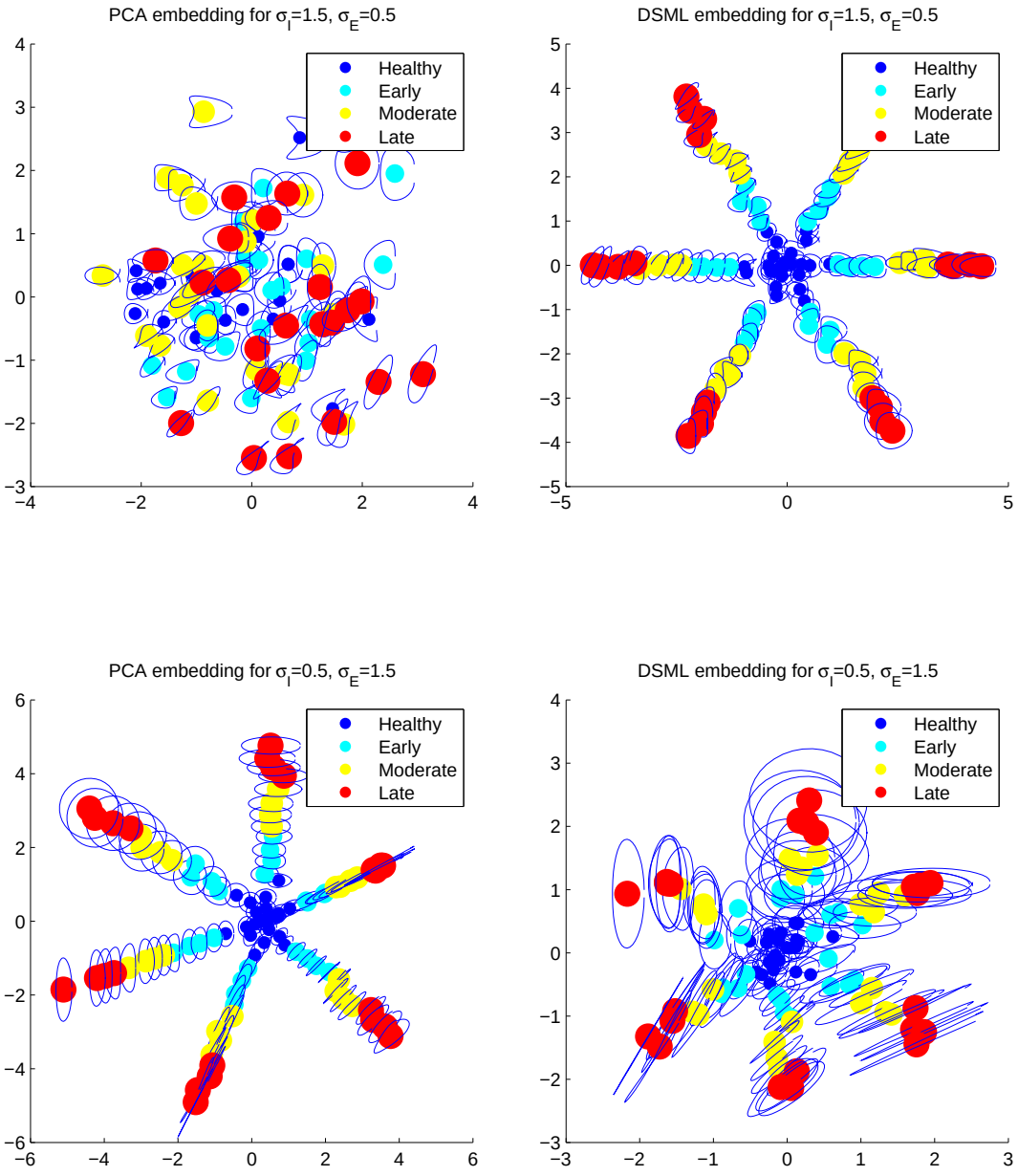


Figure 6.3: We compare the embeddings produced by the ratio of intrinsic to extrinsic variances. When the intrinsic variance is higher than the extrinsic, DSML clearly extracts the extrinsic components (3 for our synthetic data, in both positive and negative directions), while PCA does not (top). When extrinsic variance is higher than the intrinsic, PCA and DSML produce similar embeddings (bottom).

6.5.3 Classification on Synthetic Data

We now consider the classification problem: given a shape model, output a discrete stage label (e.g., healthy, early, moderate or late). We evaluate our approach for the classification by first performing dimensionality reduction and use a multiclass SVM classifier using latent variables as input. We compare our approach against: support vector machines (SVM) with raw features, SVM with features from linear discriminant analysis (LDA) embedding, SVM with features from PCA embedding and Naïve Bayes. Every algorithm is evaluated in a 3-fold cross validation setting. The total number of samples in our evaluation dataset was $N = 300$, with an equal number of samples in all 4 classes: healthy, early, moderate and late. For the multiclass SVM classifier, we used a radial basis function kernel and chose the best kernel width σ using grid sampling. We are interested in the performance of our algorithm under various intrinsic and extrinsic variances. In the first experiment, we fix the intrinsic variance $\sigma_I = 1$ and compute the misclassification rate (MCR) for all algorithms as a function of the extrinsic variance. The misclassification rate is computed as follows:

$$MCR = \frac{\text{Number of Incorrect Classifications}}{\text{Total Number of Classifications}} \tag{6.10}$$

We show the results in Figure 6.4. In the second experiment, we fix the extrinsic variance ($\sigma_E = 1$) and compute MCR for the classification algorithms performance as a function of the intrinsic variance. Results are shown in Figure 6.5.

6.6 Embedding of RA hand shapes

We now show the results of the embedding computed for our shapes extracted from radiographs of RA patients. Our (small) dataset consists of 28 healthy hands, 49 early staged, 18 moderate staged and 21 late staged hands. With the regularization term weight $\lambda = 0.02$, we show the embedding in Figure 6.6.

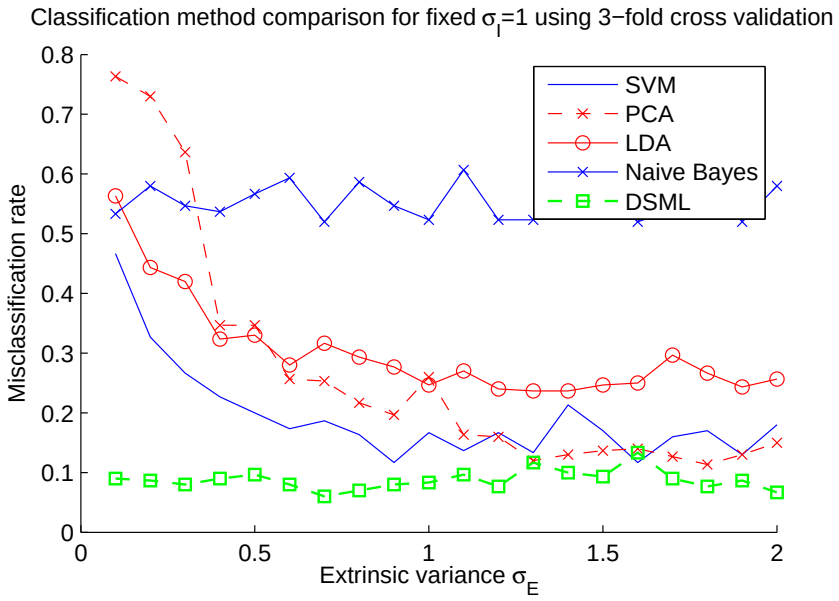


Figure 6.4: With fixed intrinsic variance ($\sigma_I = 1$), we plot the MCR for DSML, SVM PCA, LDA and Naïve Bayes classifiers. Our method outperforms all algorithms at a consistent 0.1 MCR.

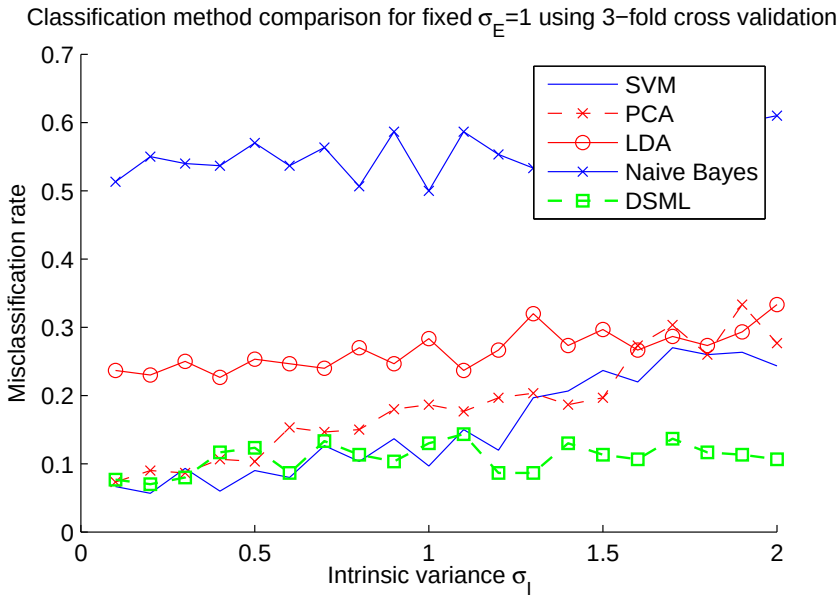


Figure 6.5: With extrinsic variance ($\sigma_I = 1$), we plot the MCR for DSML, SVM PCA, LDA and Naïve Bayes classifiers. Our method outperforms all algorithms at a consistent 0.1 MCR. Notice the consistency of our method when $\sigma_I > 1$.

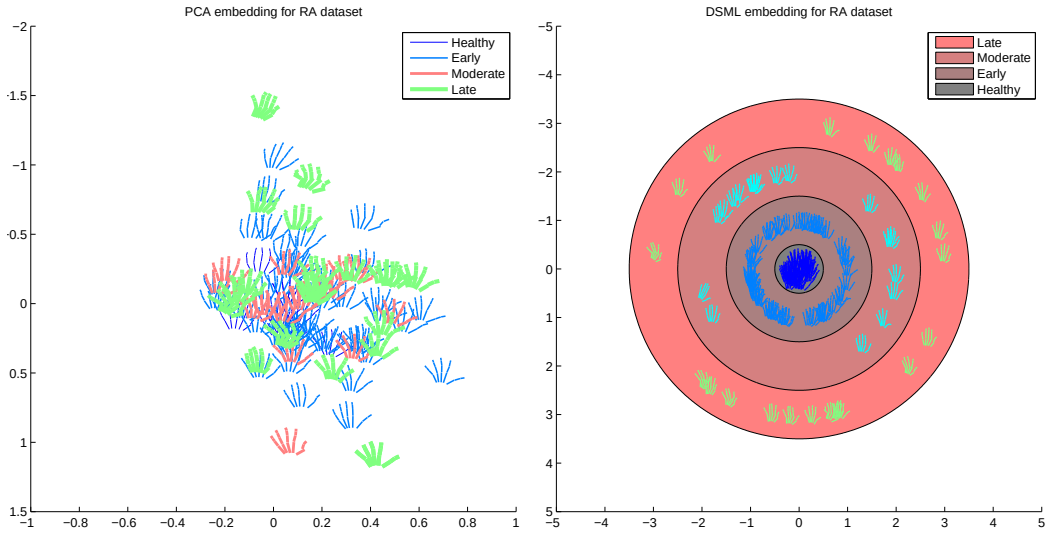


Figure 6.6: PCA embedding for RA dataset (left). DSML embedding for the RA dataset with regularization term $\lambda = 0.02$ (right).

When $rank(W) \leq 3$, we can visualize the basis vectors as a function of coefficients. This visualization provides an insight into the shape changes with respect to basis coefficients. Figure 6.7 shows such a visualization.

6.7 Implementation Details

We used an off-the-shelf optimizer from MATLAB to minimize our cost functions.

6.8 Conclusions

In this chapter we presented a novel metric learning technique for anatomical shapes. We show classification in the embedded space outperforms state-of-the-art algorithms. The intuition behind this method is that healthy shapes should map on or close to the origin of the latent space, while diseased shapes should map at a distance from the origin proportional to the stage they were labeled with.

It is a more realistic case when there are multiple shape samples from the same patient as the disease progresses. Our model can be modified to facilitate such data by including

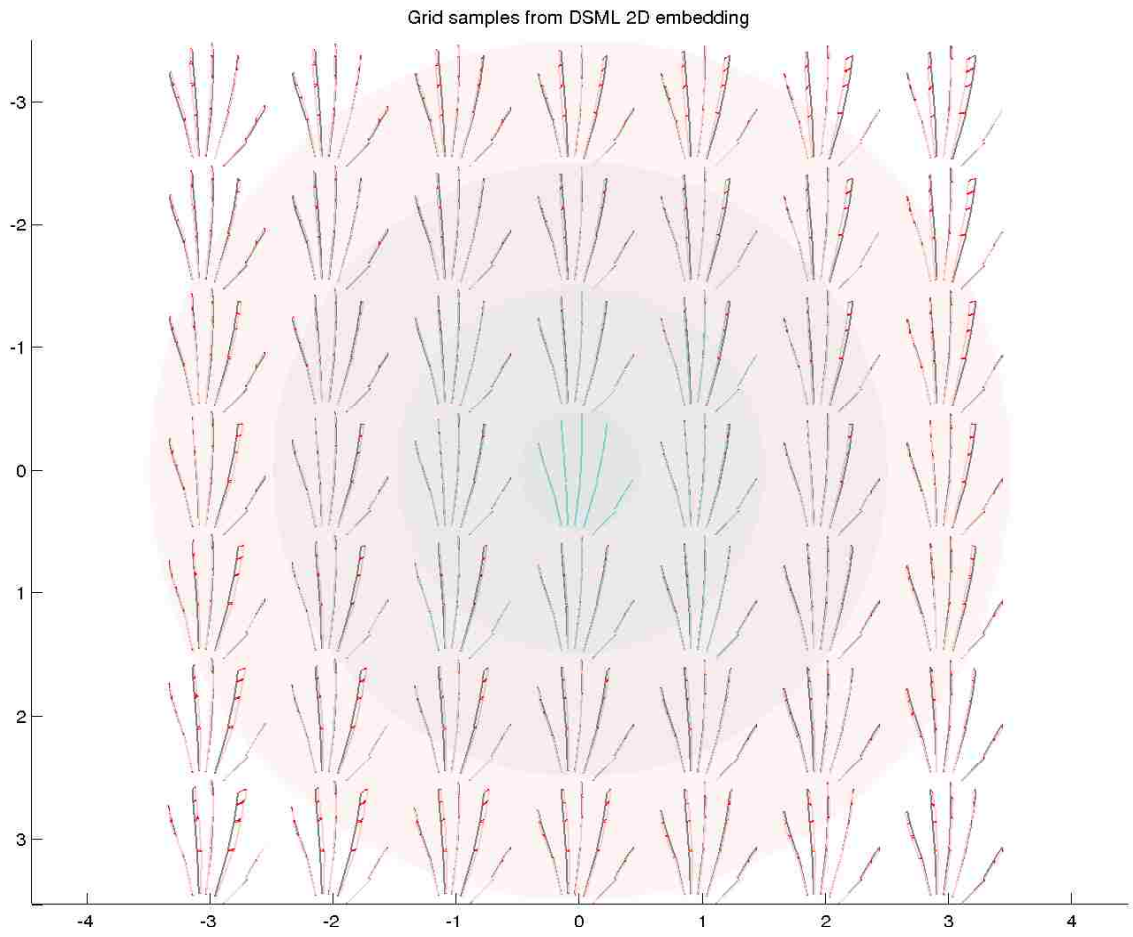


Figure 6.7: DSML embedding with grid sampling of basis coefficients when $rank(W) = 2$. The center of the embedding corresponds to the average hand. In red we show the displacements along the coefficients for the basis solved using DSML.

constraints on sequences of shapes in the embedded space. The benefit of this modification is most evident for the disease progression prediction problem: given a set of two or more shapes, what is the most likely progression path? Since the magnitude of shapes in the embedded space can be interpreted as stage, one can sample along a line from the origin to some point on the upper bound stage hypersphere.

We add a term that encourages all points from the same patient with stage labels to be close to a the average vector for the sequence in the embedding. Since each shape in the sequence has an associated stage vector, the magnitude in the embedding should still relate to stage and not conflict with the term in our original minimization problem. To this

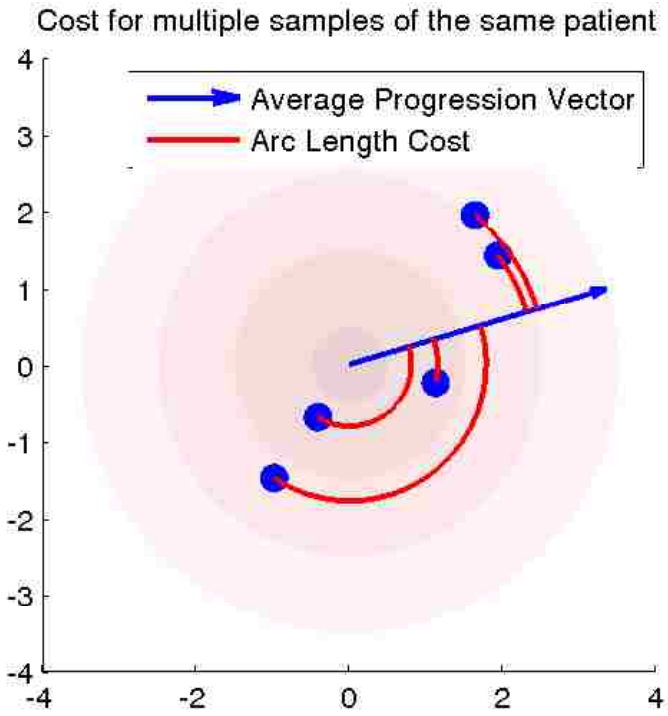


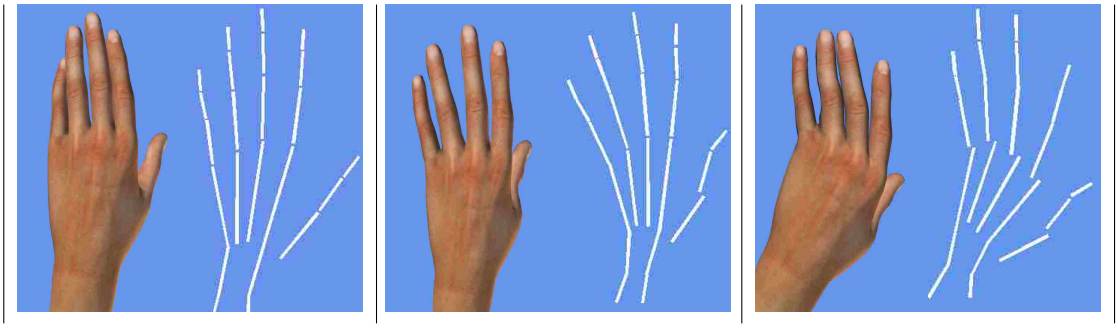
Figure 6.8: We illustrate the model term that encourages sequences from the same patient to lie on the same line. The penalty term is the sum of the shortest geodesics (arc lengths in 2D) from their current location in the embedding to the average vector.

end, our penalty term is the sum of the lengths of the shortest geodesics (arc length in 2D) from the current location to the average vector for the sequence. We illustrate this point in Figure 6.8.

This modification involves the addition of a penalty term to the cost function, term that encourages sequences of samples at varying stage to be collinear. The distance from a current location in latent space to a specific progression line is the shortest geodesic.

Chapter 7 Visualizing Hands Affected by RA

This chapter has appeared in the printed proceedings of the International Conference on Computer Games (CGAMES) 2012 [15]. In this chapter we present an approach to use high level data from hand X-Rays in postero-anterior (PA) view for synthesizing the animation and deformation of 3D hand models. We discuss our approach to model musculoskeletal deformations caused by rheumatoid arthritis (RA) and show how this could be used to animate progressions [1] of RA from early to late stage. We also discuss a potential use case of our method as a means to visualize damage due to RA as a function of time and treatment.



Early stage Moderate stage Late stage

Figure 7.1: A sample from our results. Here we show our rendered model next to the wireframe information extracted from radiographs.

7.1 Introduction

Serious games serve purposes other than pure entertainment. They can be used in a wide range of applications, from assisting people in recovering from trauma to persuading patients with early signs of debilitating diseases to take their treatment seriously. In all cases, the success of a serious game is largely affected by how realistically it can depict the aspects

[1] RA patients often experience different types of deformities, occurring at unpredictable times after the onset of the disease.

of the simulation that it is trying to emphasize. As part of a serious game, the work in this paper uses a data-driven approach from radiographs (X-rays) to demonstrate the amount (and various types) of deformations RA patients could experience. Using sample X-Rays at different stages of RA, we animate a rigged model of a hand to show the progression of the disease.

In this chapter, we will explain how high level data from X-Rays can be used to better model deformations of skeletal structures afflicted with RA in a computer generated animation.

7.2 Background

RA is a chronic, autoimmune disease that leads to swelling of synovial fluid around the joints[31]. This fluid builds up and can cause joint erosion, fusing of the joints, and could potentially lead to other problems with other organs in the body. Our work focuses on the skeletal changes in the hand during the progression of this disease. While other 3D models of hands have been created for the study of hand dexterity [5], our work shows the skeletal deformations that occur as the disease progresses.

The work described here is part of a larger project to design and build a decision aid for patients at the onset of this disease. Treatments are available to suppress some of the symptoms and help ease the pain, but it is up to the patients whether or not they will adhere to these treatments. In a recent study [78], use of decision aids led to a better understanding of the disease for the patients. Simulations are one of the most prominent applications for hand modeling in computer graphics today [5], however, serious games used as a decision aid for patients are scarce for topics other than mental health or fitness [74]. This genre of gaming can allow users to experience situations that would be impossible in the real world [98]. Our work is a contribution to the development of serious games, and could generalize to other musculoskeletal disorders with no cure. This work allows patients to see what is most likely to happen to their hand if no treatment is used or in the case of poor adherence

to a prescribed treatment.

The data we used in our work is a series of X-Rays of different patients in three stages of RA, categorized by an expert as early, moderate, and severe. The user can select a hand from each category; our approach then interpolates between the 3 hand stages and animates the progression, from early to severe. The user sees a 3D model on screen, showing the resulting deformations that occur.

Our work could be used by developers of medically oriented serious games to depict realistic bone deformations in their characters. Our approach uses a skinned hand mesh that is deformed according to information extracted from X-rays and can easily be attached to a character.

Human hands are complex anatomical structures able to perform fine motor tasks as well as powerful grasping. We need them for day to day activities. Most people take them for granted until disaster strikes (i.e., they lose them or they get diseased). For RA patients, losing dexterity due to pain and deformities is reality. In this chapter we present a method based on character animation to inform patients about the disease and give them a glimpse of a possible future given a choice in their health care. In order to make an impact in the patients' perception of the disease and how treatment choices will impact them, we propose to use a hand model to reflect deformities caused by RA. In order to do that, we dive into computer graphics and look at existing work in anatomically correct hand modeling and apply what we learn from radiographs to deform the hands correctly. This is not a trivial task due to several considerations: anatomical complexity of the human hand, specific constraints on motion due to the pain and deformations and computational demands. The field of biomechanics looks at anatomical structures with the objective to understand intricacies of motions in living things. The primary goal of this component of the project is to create believable **visualizations** of deformed hands. In the future, we plan to use this work as building blocks for the analysis of biomechanical limitations RA patients suffer in any stage of the disease.

Character Animation Granberg [39] provides an excellent survey of the history of character animation. In the early days of computer games, characters have been small grids of colored pixels that were animated by making multiple pictures with the character in a different pose. Nowadays, graphics cards are capable of performing astounding amounts of computation, allowing for more complex animation schemes. 3D models are typically stored as meshes consisting of connected vertices. There are two popular methods to animate 3D models: morphing animation and skeletal animation. In morphing animation, two or more meshes are blended on a per-vertex basis [39]. Consequently, the meshes have to have the same number of vertices and polygons need to be arranged similarly in order for the technique to work. Historically, morphing animation has been used for full character animation. Its main drawback is it requires considerable artist effort. Currently, morphing animation is used for facial animation [39]. Another, more popular approach to animating 3D models is skeletal animation. Skeletal animation works by introducing invisible "bones" (in practice, these are just transformation matrices in a hierarchy) in a mesh where the position of the bones determines a pose. As bones move, the mesh around them deforms. The deformation of the mesh with respect to the bone positions is called **skinning**. Skinning is an open research area, with the aim to improve the visual quality of the deformed mesh. The most popular skinning technique is linear blend skinning also known as Skeleton Subspace Deformation (SSD) introduced by Magenat-Thlamann et al. [60]. The main idea is that bones are transformed rigidly, while vertices in the mesh are deformed by a linear combination of the neighboring bones transformations. In other words, one vertex can be transformed by multiple bones determined by weights. This method relies on an artist and modeling software that "paint" weights on vertices in order to create realistic deformations. SSD is a simple yet effective skinning technique, but it has shortcomings (e.g., "candy-wrapper" and "collapsed elbow" effects). These problems become visible when complex articulated objects are deformed, causing them to appear unnatural. In this research, our aim is to render and animate human hands, which are highly articulated and

complex structures. Moreover, we want to render and animate hand models with deformities caused by RA. In order to realize our proposition, we require an anatomically corect skeletal model that is deformed according to a progression scale and is animateable. Given the complexity of the task and the known shortcomings of SSD, we need a novel skinning technique that is both aesthetically pleasing and conforms to anatomical constraints. In the following paragraph we provide a brief survey of existing work.

Graphics for Hand Modeling We now turn our attention to existing in-silica models of human hands. For the larger scope of this project, we hypothesize that a realistically animated avatar hand will be taken more seriously by patients, thus improving the impact of our proposed intervention. In order to realize that, we survey existing work and propose a solution based on previous research and suggest other research questions that can be addressed.

Skeletal poses are determined by the amount of deviation from a rest pose caused by skeletal muscles. There are two widely used mathematical models that describe functional and mechanical properties of muscles: one originating from Hill [44] and another from Huxley [47]. Hill-based models consist of elastic elements in parallel with a series elastic element and a contractile element. Huxley-based models are used to understand properties of microscopic contractile elements [99].

The contraction of muscles alters bone positions and also produces visible effects (e.g., bulging) to the skin. In this work we are interested in both forces and effects. A survey by Lee et al. [57] summarizes developments in the area of muscle simulation in computer graphics. They describe the anatomy of skeletal muscles as well as previous experiments that consist of determining force exerted by different types of muscles. Our primary research focus is to use artificial muscle models to drive skeletal deformation, in turn producing realistic deformations. While the biomechanics aspect of RA deformed hands poses interesting research questions, we will focus on visualizing RA effects rather than a metic-

ulous analysis of force constraints posed on the RA affected human hand, which we plan to investigate in the future.

Computational approaches to muscle simulation can be categorized in three classes [57]: geometrically based, physically based and data driven. Geometrically based approches focus on modeling animation effects of muscle contraction. They have been shown to be successful in modeling simple muscles, but are not straightforwardly extended to more complex muscles [57]. Such examples are Free Form Deformations [17] and parametric and polygonal surfaces [106, 100].

Sueda et al. [96] tackle the realism of hand rendering problem by incorporating tendons and muscles under the skin of traditionally animated characters. Their method relies on two primitives: rigid bodies (bones) and strands based on cubic B-spline curves. The idea is that the motion of the spline control points is spatially constrained in order to route tendons that deform the skin above them and also drive the skeleton transformations. They employ two types of constraints for tendon routing: sliding and surface constraints. The sliding constraints are used when a strand has to pass through a specific point in space (near a bone) while a surface constraint allows for lateral strand movement as well. Fixed constraints are used for strand origins and insertions. The authors developed a controller to determine activations of each tendon (equivalent of muscle forces and velocities) in order to deform the model to a certain pose. Their main contribution is the computation of dynamic simulations where complex routing constraints are required. In terms of efficiency, they report the computation time required to get the activation levels of tendon strands for a few seconds taking only a few minutes. While this is impractical for a real time implementation, they do not mention a possible GPU implementation, which could be investigated.

Data driven approaches directly model the skin shape based on observations from range scans [6], silhouettes from video streams [90] and other sources such as CT scans [54]. Kurihara et al. [54] propose a method for building hand models from CT (Computed Tomography) scan data. Their method uses volumetric CT data that is first segmented

into bone tissue and skin. They then estimate the joint rotation centers and joint angles of several scans. An arbitrarily chosen base model is then deformed to a desired pose using Weighted Pose Space Deformation (a combination of SSD and morphing) and then fitted interactively to meshes of other scans using feature points. Data driven approaches are the least appealing for our project due to lack of motion data. Physically based approaches present the most potential for this project due to the underlying goal of modeling muscle contraction forces and representing changing muscle geometry during contraction [57].

Physically based approaches rely on physics to tackle muscle dynamics and tissue properties. This problem requires addressing the problems of how to determine contractile muscle forces and how to represent the changing muscle geometry during contraction. Several computational models have been proposed: mass-spring systems, Finite Element Method (FEM) and Finite Volume Method (FVM) [57]. In mass-spring systems, objects are modeled by a collection of points linked by massless springs. The deformation effects are computed using energy minimization techniques that compute equilibria based on **changing** spring properties during contraction. This model can also be extended to include other types of spring forces (e.g., angular, bending and shearing) [57]. Albrecht et al. [5] propose hand models based on an underlying anatomical model (bones) that is animated using muscle contraction values. Their contribution is a hybrid muscle model comprising of pseudo-muscles and geometric muscles. The pseudo muscles control bone rotation according to anatomical constraints and mechanical laws, while the geometric muscle cause realistic bulging of skin tissue. The geometric muscles have a geometrical shape that deforms and bulges during contraction. They connect the geometric muscles to the skin and bones using a mass spring system, inspired by [49]. Their method requires a significant effort in tuning parameters which is undesirable and also lacks models of tendons which should drive the skeleton movements using geometric muscles instead of the pseudo muscles.

In the Finite Element Method (FEM) approach, a body (e.g., muscle) is subdivided into discrete elements (hexahedra or tetrahedra in 3D) where the displacements and positions of

an element are computed using an interpolation function. In a dynamic problem, equilibria is computed by solving partial differential equations (PDEs) using FEM. Tang et al. [99] developed a 3D computational model based on FEM to describe the active and passive non-linear mechanical behavior of skeletal muscles. They report their numerical algorithm is capable of determining shortening and lengthening of muscles due to concentric and eccentric contraction. Finite Volume Method (FVM) is similar to FEM in that PDEs are modeled by algebraic equations [57].

7.3 Development

7.3.1 Data Acquisition

The data used in the decision aid was acquired from a set of anonymous X-Rays taken from patients with various stages of RA. We model the deformations using a wireframe representation of the hand, similar to that of Martín-Fernández et al. [66] (see Figure 7.2), but with an important modification: we allow the segments to be disconnected. This modification allows subluxation (dislocation of joints, common in RA) to be modeled.

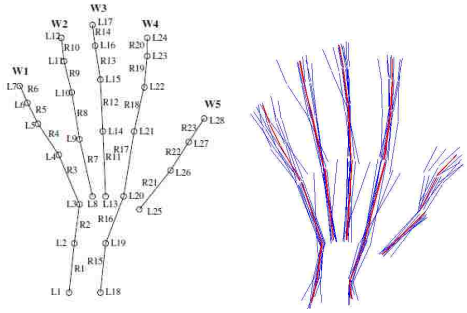


Figure 7.2: Left: Wireframe representation of Martín-Fernández et al. There are 5 wires, with each segment positioned on landmarks in the X-rays. Right: our data consisting of manually clicked points on radiographs after alignment (we show the average hand in red).

We manually annotated 18 radiographs (6 in each stage) by clicking on landmarks that correspond to the wireframe model (shape). Due to inconsistent resolution and calibration parameters of the X-ray machines, we performed an alignment step using Generalized

Procrustes Analysis (GPA). This method computes an optimal rigid body transformation for each shape that minimizes a global similarity metric. Our data after the alignment step can be seen in Figure 7.2.

Once the shapes are aligned, we use their skeletal hierarchy to get relative rotations and displacements of the bones and apply those transformations locally to our rigged mesh model. The set of transformations given by a single hand data is used as a keyframe in the animation of the hand. As the hand interpolates between keyframes, the viewer is able to see the progression of RA on a synthetic hand.

7.3.2 Model

We used a hand model rigged with bones that match our wireframe representation. This abstraction (while not entirely anatomically accurate, e.g., carpal bones are ignored) allows plausible realistic animations. Since bones in the mesh model match the wireframe representation, we can compute the transformations from the data (rotations and translations) for rendering. Different textures (e.g., varying skin color, adding wrinkles, etc.) can be used to customize the model, so that patients can better relate to the avatar in our proposed application.

Figure 7.3: Model used in the simulation.

We generate the keyframes for animation by computing the rotations and translations in the wireframe model and apply them to our rigged model using forward kinematics. We

show the results in the next section. It is important to mention that plain radiographs are projection images, thus volumetric information is lost. The hand is a complex biomechanical structure with joints of varying degrees-of-freedom (DOF). In this work we rely on the pose constraint of the hand during image formation, namely the PA view that enforces a known kinematic configuration. Deformities due to RA violate these constraints and are useful for visualization and further investigation in anatomically correct musculoskeletal anomaly modeling.

7.4 Results

As shown in Fig. 1, we achieved convincing results using this technique.
Figure 7.1a shows a deformed hand from one of the early stages in our sample data. One can see from this image that the knuckles are not visibly affected compared to the base model shown in Figure 7.3. The main difference is the orientation of the metacarpal bones as they have slight deviation from the base model (here considered healthy).

Figure 7.1b shows a hand with moderate RA. It has noticeably more displacement in the knuckles as well as more severe deviation of the metacarpal bones. Figure 7.1c shows a hand with severe RA. One can see that the shape of the model is significantly disfigured compared to the base model. The knuckles show more displacement and ulnar deviation is visibly more severe from the way they were in the base model. Figure 7.4 shows photographs of 3 different people who are suffering from RA. From left to right we see the progression from early to severe.

7.5 Future Work

Our approach, while visually convincing and useful as a visualization component of a decision aid (DA) to help patients become actively involved in their care, can be improved through the use of volumetric (CT or MRI) data. In current RA care, it is less common for CT or MRI scans to be ordered, as opposed to the more common (and inexpensive) plain

Figure 7.4: Examples of severe RA hand deformations. Retrieved from [3].

radiographs. Due to the commonality and availability of plain radiographs, our decision tool can be customized for the patient, by automatically extracting the wireframe landmarks from a recent radiograph, and showing possible disease progressions from the status-quo, contingent on treatment choice.

One of the first signs of RA visible on a radiograph is swelling of the joints. We are currently working on a method to automatically detect the amount of swelling from radiographs and apply them to our model.

Given a large set of hand radiographs from early to severe, it is possible to use machine learning techniques to discriminate between specific types of deformations common to RA. An example application of such results would be to accurately predict the course of the disease for a specific patient given a sequence of radiographs and conditioning on treatment. The larger project of constructing a DA for patients to decide on a treatment is based on the evidence from the medical literature that chronic illness patients, such as those with RA, often do not adhere to the prescribed treatment [52, 79]. This is in part due to a long onset time of drugs and the (early) presence of side-effects. We conjecture that patients who use this tool can avoid the pitfalls associated with poor adherence.

In the future we intend to add animations of the hand interacting with objects, and study the limitations caused by RA from the biomechanics literature and make contributions to the anatomically accurate modeling literature as well as computer graphics and DA design.

7.6 Conclusion

Our work is a step toward anatomically and medically accurate models of human characters for serious games. Using a data-driven approach, we were able to successfully model the progression of rheumatoid arthritis in a model of a hand to demonstrate the damaging effects of the disease. This approach shows how the animation and modeling of diseases can be used to improve the graphical realism of a serious game. Using this approach, serious games can be developed to incorporate information from medical imaging data collected to increase realism of animations. While we used RA as a sample disease, this approach can be generalized to other musculoskeletal diseases for which medical image data is available and a means to annotate them exists.

Chapter 8 Conclusions and Future Directions

In this dissertation, we presented work on several computational problems that are essential to enhance the status quo of patient-centered health care. In particular, we took the perspective of patients who are newly diagnosed with diseases that have uncertain progressions and the treatments have uncertain benefits and side-effects.

Health care economics can severely limit the time a patient spends with a specialist. For an average person to understand complex diseases and treatments, the medical practitioners have to adapt their explanations to a wide gamut of patient literacy and numeracy. Computational tools that facilitate this process and mold a more informed and empowered patient are critical for improving health care. This dissertation uses rheumatoid arthritis as the choice of disease because it has several characteristics shared by other chronic diseases, which one can apply the work in this dissertation.

In Chapter 3 we take a comprehensive look at patient decision aids and focus in on work done specifically for RA patients. Unfortunately, few DAs are available for RA at the time of this writing and the existing DAs lack in key areas such as risk communication, evidence citing and customization options. We argue that developers of DAs should aim to make the products well structured, tailored and/or interactive. Much of the work in this dissertation is geared to provide DA developers with the computational tools needed to make their products interactive and customizable. While much of the focus is on RA, our findings generalize to other chronic diseases.

In Chapter 4 we propose a method for patients with deformities of the hands (e.g., caused by RA) to interact with a computer via gestures. The basic idea is that our system can easily be set up in a doctor's office where an assistant can help the patients select (or even create) a set of gestures that are possible for their deformity stage. Our system makes use of two Microsoft Kinect sensors that provide depth images in real time out from which

we can train the system to recognize a set of gestures that RA patients can perform and use them for activities that can range from simple selection operations to more complex interactions such as navigating a 3D world. Our method requires modest hardware and is highly customizable.

In Chapter 5 we propose a method to automatically extract shape information from hand radiographs in postero-anterior (PA) view. We created a point distribution model (PDM) that is flexible and able to capture key deformities caused by RA: joint space narrowing and subluxation (shifting of bones). This problem is challenging due to the wide variation in appearance and pose in radiographs, compound with the destructive effects of RA. Our method uses partial probabilistic information from randomized decision forests trained with dense SIFT features. We formulate the shape fitting problem as inference in a conditional random field framework. We design potential functions tuned for specific anatomical structures and a learned shape prior. We evaluate our approach on a publicly available dataset of radiographs of relatively healthy hands and one of RA patients at different stages. The output of this method can be used to build models of disease progression (as we show in Chapter 6). We also show an alternative use as an initialization method for algorithms to automatically estimate bone contours, which is a challenging problem and important in assessing disease progression.

In Chapter 7 we exploit the structure of our PDM and use samples to deform a 3D hand model. This method provides novel visualization of the effects of RA on hands. The appearance of the hand model can be customized and animations of progression are data driven. This approach is flexible and patients can receive visual answers to the possible appearance of their hand in the future.

In Chapter 6 we propose an algorithm that learns a distance metric for use on anatomical shapes. The intuition is that anatomical shape variability can be divided in two groups: intrinsic variability and extrinsic variability. Intrinsic variability stems from genetic differences between subjects (e.g., bigger hands), while extrinsic variability is solely due to

a disease process. A good distance metric should return low to zero between any shapes considered healthy (despite drastic differences in appearance) and high between healthy shapes and shapes that are staged as advanced by a medical professional. The distance metric can be used to create a latent embedding where the goal is to "collapse" all healthy shapes to the origin while diseased shapes are at a distance from the origin proportional to the stage.

This dissertation brings contributions in several areas of computer science. We feel that data driven, interactive and customizable patient decision aids are a key for better health care. Empowered and informed patients are better patients. The computational tools we presented in this dissertation are a good start for data-driven interactive decision aids. We think that giving patients glimpses of possible futures is a powerful tool and motivator to be active participants in their health care.

Big data is already here and research efforts are underway to make sense of it. With inexpensive digital storage, large amounts of medical imagery are now available. We believe that "big medical data" is approaching and computer vision can play an important role to increase our understanding of the disease processes visible through medical imaging modalities. With better insights into disease specific changes, we can build better pictures of possible futures for patients who have to make difficult decisions.

It takes years of training for specialists to stage a disease from medical imagery. Automated methods can not only help reduce human errors, but they can paint a better picture of the different modes of variation. Tools from probability theory can be used in tandem to reason in light of uncertainty. One route that we are planning to pursue is an extension of the work in Chapter 6 to include treatment information. Predicting the course of the disease for a patient *given a certain drug* can have a powerful effect on the decision making process.

Visualizing possible RA disease progression paths using virtual models deformed using SSD as described in Chapter 7 can be improved to include swelling. We plan on extending

the work in Chapter 5 to detect swelling from plain radiographs and improve the anatomical accuracy and realism of the 3D models.

Radiographs are projections of 3D objects, hence volumetric information is lost. Given an anatomically accurate 3D hand model, we will try to solve for the (approximate) kinematic configuration of the hand for a given radiograph by simulating X-rays. GPUs are becoming inexpensive and powerful general purpose computing devices; we can harness their power for X-ray simulations and recover more information from a flat radiograph than previously possibly. Deformities make this problem more challenging, but the results would be useful in both visualization and medical image processing realm.

Bibliography

[1] International Patient Decision Aid Standards (IPDAS) Collaboration. http://ipdas.ohri.ca/index.html.

[2] Patient education - diseases and conditions. http://www.rheumatology.org/practice/clinical/patients/diseases_and_conditions/index.asp.

[3] Arthrite rheumatoide. http://upload.wikimedia.org/wikipedia/commons/3/3a/Arthrite_rhumatoide.jpg, 2006. Accessed May 2013.

[4] Santiago Aja-Fernández, Rodrigo de Luis-Garcıa, Miguel Angel Martın-Fernández, and Carlos Alberola-López. A computational TW3 classifier for skeletal maturity assessment. a computing with words approach. *Journal of Biomedical Informatics*, 37(2):99–107, 2004.

[5] Irene Albrecht, Jörg Haber, and Hans-Peter Seidel. Construction and animation of anatomically based human hand models. In *Proceedings of the 2003 ACM SIGGRAPH/Eurographics symposium on Computer animation*, pages 98–109. Eurographics Association, 2003.

[6] B. Allen, B. Curless, and Z. Popović. Articulated body deformation from range scan data. In *ACM Transactions on Graphics (TOG)*, volume 21, pages 612–619. ACM, 2002.

[7] Vassilis Athitsos, Haijing Wang, and Alexandra Stefan. A database-based framework for gesture recognition. *Personal and Ubiquitous Computing*, 14(6):511–526, 2010.

[8] L. Ballerini and L. Bocchi. Multiple genetic snakes for bone segmentation. *Applications of Evolutionary Computing*, pages 346–356, 2003.

[9] David J. Bartholomew, Martin Knott, and Irini Moustaki. *Latent variable models and factor analysis: a unified approach*, volume 899. John Wiley & Sons, 2011.

[10] N.D. Berkman, Sheridan S.L., K.E. Donahue, D.J. Halpern, A. Viera, K. Crotty, A. Holland, M. Brasure, K.N. Lohr, E. Harden, and et al. *Health literacy interventions and outcomes: An updated systematic review*. Agency for Healthcare Research and Quality, 2011.

[11] J.P. Berrios-Rivera, Jr. R.L. Street, M.G. Garcia, M.A. Kallen, M.N. Richardson, N.M. Janssen, D.M. Marcua, J.D. Reveille, N.B. Warner, and M.E. Suarez-Almazor. Trust in physicians and elements of the medical interaction in patients with rheumatoid arthritis and systemic lupus erythematosus. *Arthritis Rheum*, 55(3):385–393, 2006.

[12] Z. Zenn Bien, Kwang-Hyun Park, Jin-Woo Jung, and Jun-Hyeong Do. Intention reading is essential in human-friendly interfaces for the elderly and the handicapped. *Industrial Electronics, IEEE Transactions on*, 52(6):1500–1505, 2005.

[13] Maya Buch and Paul Emery. The aetiology and pathogenesis of rheumatoid arthritis. *Hospital Pharmacist-London*, 9(1):5–10, 2002.

[14] Nicholas Burrus homepage - kinect calibration. `http://nicolas.burrus.name/index.php/Research/KinectCalibration`, November 2010.

[15] Kaitlin Burton, Frederick Hallock, and R. Paul Mihail. A data-driven approach to visualizing the effects of rheumatoid arthritis. In *International Conference on Computer Games (CGAMES)*, 2013.

[16] Eric Cassell. *The healer's art*. MIT Press, 1985.

[17] J.E. Chadwick, D.R. Haumann, and R.E. Parent. Layered construction for deformable animated characters. In *ACM Siggraph Computer Graphics*, volume 23, pages 243–252. ACM, 1989.

[18] H.Y. Chai, L.K. Wee, T.T. Swee, and S. Hussain. Glcm based adaptive crossed reconstructed (acr) k-mean clustering hand bone segmentation. *Recent Researches in Communications, Automation, Signal Processing, Nanotechnology, Astronomy and Nuclear Physics*, pages 192–197, 2011.

[19] Cathy Charles, Amiram Gafni, and Tim Whelan. Shared decision-making in the medical encounter: what does it mean?(or it takes at least two to tango). *Social science & medicine*, 44(5):681–692, 1997.

[20] F. Chilton and R.A. Collett. Treatment choices, preferences and decision-making by patients with rheumatoid arthritis. *Musculoskeletal Care*, 6(1):1–14, 2008.

[21] Tim F. Cootes, Mircea C. Ionita, Claudia Lindner, and Patrick Sauer. Robust and accurate shape model fitting using random forest regression voting. In *Computer Vision–ECCV 2012*, pages 278–291. Springer, 2012.

[22] David Cristinacce and Timothy F. Cootes. Feature detection and tracking with constrained local models. In *BMVC*, volume 17, pages 929–938, 2006.

[23] L.H. Daltroy. Doctor-patient communication in rheumatological disorders. 7(2):221–239, June 1993.

[24] Luke Davis, Barry-John Theobald, Andoni Toms, and Anthony Bagnall. On the extraction and classification of hand outlines. In *Intelligent Data Engineering and Automated Learning-IDEAL 2011*, pages 92–99. Springer, 2011.

[25] Kevin D. Deane, Jill M. Norris, and V. Michael Holers. Pre-clinical rheumatoid arthritis: Identification, evaluation and future directions for investigation. *Rheumatic diseases clinics of North America*, 36(2):213, 2010.

[26] G. Elwyn, A.M. O'Connor, Stacey D., R. Volk, A. Edwards, A. Coulter, R. Thomson, A. Barratt, M. Barry, Bernstein S, et al. Ipdas 2005: Criteria for judging the quality of patient decision aids, 2005.

[27] Ali Erol, George Bebis, Mircea Nicolescu, Richard D. Boyle, and Xander Twombly. Vision-based hand pose estimation: A review. *Computer Vision and Image Understanding*, 108(1-2):52–73, 2007.

[28] A. Fagerlin, P.A. Ubel, and C. Wang. Reducing the influence of anecdotal reasoning on peoples health care decisions: is a picture worth a thousand statistics? *Med Decis Making*, 25:398–405, 2005.

[29] D. Feldman-Stewart, S. Brennenstuhl, K. McIssac, Austoker J., Charvet A., Hewitson P., Sepucha K.R., and Whelan T. A systematic review of information in decision aids. *Health Expecations*, 10(1):46–61, 2008.

[30] Ronald A. Fisher. The use of multiple measurements in taxonomic problems. *Annals of eugenics*, 7(2):179–188, 1936.

[31] Arthritis Foundation. Rheumatoid arthritis. `http://www.arthritis.org/conditions-treatments/disease-center/rheumatoid-arthritis/`. Accessed May 2013.

[32] L. Fraenkel, Bogardua S, Concato J, and Felson D. Risk communication in rheumatoid arthritis. *Journal of Rheumatology*, (30):443–448, 2003.

[33] Bláithín Gallagher and Helen Petrie. Initial results from a critical review of research on technology for older and disabled people. In *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility*, page 53. ACM, 2013.

[34] G. Gigerenzer. The psychology of good judgment: Frequency formats and simple algorithms. *Medical Decision Making*, 3(16):273–280, July 1996.

[35] G. Gigerenzer and A. Edwards. Simple tools for understanding risks: from innumeracy to insight. *BMJ*, 327(7417):741–744, 2003.

[36] Daniela Giordano, Concetto Spampinato, Giacomo Scarciofalo, and Rosalia Leonardi. An automatic system for skeletal bone age measurement by robust processing of carpal and epiphysial/metaphysial bones. *Instrumentation and Measurement, IEEE Transactions on*, 59(10):2539–2553, 2010.

[37] Amir Globerson and Sam Roweis. Metric learning by collapsing classes. In *Nips*, volume 18, pages 451–458, 2005.

[38] Jacob Goldberger, Sam Roweis, and Geoffrey Hinton. Neighborhood components analysis. In *Advances in Neural Information Processing Systems*, volume 17, pages 513–520, 2005.

[39] Carl Granberg. *Character Animation with Direct3D*. Charles River Media, 2009.

[40] M.L. Grove, A.B. Hassell, and Hay EM et al. Adverse reactions to disease-modifying anti-rheumatic drugs in clinical practice. *Q J Med*, 94(6):309–319, 2001.

[41] Hans Werner Guesgen and Darren Kessell. Gestural control of household appliances for the physically impaired. In *FLAIRS Conference*, 2012.

[42] H. Hasan and S. Abdul-Kareem. Static hand gesture recognition using neural networks. *Artificial Intelligence Review*, pages 1–35, 2012.

[43] L. Haugli, E. Strand, and A. Finset. How do patients with rheumatic disease experience their relationship with their doctors? a qualitative study of experiences of stress and support in the doctor-patient relationship. 52(2):169–174, February 2004.

[44] A.V. Hill. The heat of shortening and the dynamic constants of muscle. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 126(843):136–195, 1938.

[45] I. Hirono, H. Hideki, and Y. Eiji. Patients preferences for decision making and the feeling of being understood in the medical encounter among patients with rheumatoid arthritis. *Arthritis & Rheumatism*, 55(6):878–883, 2006.

[46] T.T.M. Hue, J.Y. Kim, and M. Fahriddin. Hand bone radiograph image segmentation with roi merging. In *Proceedings of the 13th IASME/WSEAS international conference on Mathematical Methods and Computational Techniques in Electrical Engineering conference on Applied Computing*, pages 147–154. World Scientific and Engineering Academy and Society (WSEAS), 2011.

[47] A.F. Huxley et al. Muscle structure and theories of contraction. *Progress in biophysics and biophysical chemistry*, 7:255, 1957.

[48] K. Kaarela and H. Kautiainen. Continuous progression of radiological destruction in seropositive rheumatoid arthritis. *The Journal of rheumatology*, 24(7):1285–1287, 1997.

[49] K. Kahler, J. Haber, and H.P. Seidel. Geometry-based muscle modeling for facial animation. In *Graphics Interface*, pages 37–46, 2001.

[50] Michael Kass, Andrew Witkin, and Demetri Terzopoulos. Snakes: Active contour models. *International journal of computer vision*, 1(4):321–331, 1988.

[51] Daphne Koller and Nir Friedman. *Probabilistic graphical models: principles and techniques*. MIT press, 2009.

[52] K.P. Krueger, B.A. Berger, and B. Felkey. Medication adherence and persistence: a comprehensive review. *Advances in therapy*, 22(4):313–356, 2005.

[53] Sanjiv Kumar and Martial Hebert. Discriminative random fields: A discriminative framework for contextual interaction in classification. In *IEEE International Conference on Computer Vision (ICCV)*, 2003.

[54] T. Kurihara and N. Miyata. Modeling deformable human hands from medical images. In *Proceedings of the 2004 ACM SIGGRAPH/Eurographics symposium on Computer animation*, pages 355–363. Eurographics Association, 2004.

[55] Georg Langs, Philipp Peloschek, and Horst Bischof. ASM driven snakes in rheumatoid arthritis assessment. In *Image Analysis*, pages 454–461. Springer, 2003.

[56] A. Larsen, K. Dale, M. Eek, et al. Radiographic evaluation of rheumatoid arthritis and related conditions by standard reference films. *Acta radiologica: diagnosis*, 18(4):481, 1977.

[57] D. Lee, M. Glueck, A. Khan, E. Fiume, and Ken Jackson. A survey of modeling and simulation of skeletal muscle. *ACM Transactions on Graphics*, 28(4), 2010.

[58] Zhenguo Li, Jianzhuang Liu, and Xiaoou Tang. Pairwise constraint propagation by semidefinite programming for semi-supervised classification. In *Proceedings of the 25th international conference on Machine learning*, pages 576–583. ACM, 2008.

[59] Shan Lu, Dimitris N. Metaxas, Dimitris Samaras, and John Oliensis. Using multiple cues for hand tracking and model refinement. In *CVPR (2)*, pages 443–450, 2003.

[60] N. Magnenat-Thalmann and D. Thalmann. Human body deformations using joint-dependent local operators and finite-element theory. *Making them move: mechanics, control, and animation of articulated figures*, pages 243–262, 1991.

[61] Sasan Mahmoodi, Bayan S. Sharif, E. Graeme Chester, John P. Owen, and Richard Lee. Skeletal growth estimation using radiographic image processing and analysis. *Information Technology in Biomedicine, IEEE Transactions on*, 4(4):292–297, 2000.

[62] Sasan Mahmoodi, B.S. Sharif, E.G. Chester, J.P. Owen, and R.E.J. Lee. Automated vision system for skeletal age assessment using knowledge based techniques. In *IEEE International Conference on Image Processing and Its Applications*, 1997.

[63] M.N. Mamatha and S. Ramachandran. Automatic eyewinks interpretation system using face orientation recognition for human-machine interface. *IJCSNS International Journal of Computer Science and Network Security*, 9(5):155–163, 2009.

[64] M. Maravic, C. Berge, J.P. Daures, M.C. Boissier, et al. Practices for managing a flare of long-standing rheumatoid arthritis: survey among french rheumatologists. *Clin Exp Rheumatol*, 23:36–42, 2005.

[65] Mary Margaretten, Laura Julian, Patricia Katz, and Edward Yelin. Depression in patients with rheumatoid arthritis: description, causes and mechanisms. *International journal of clinical rheumatology*, 6(6):617–623, 2011.

[66] M.Á. Martín-Fernández, R. Cárdenes, E. Muñoz-Moreno, R. de Luis-García, M. Martín-Fernández, and C. Alberola-López. Automatic articulated registration of hand radiographs. *Image and Vision Computing*, 27(8):1207–1222, 2009.

[67] Iain B. McInnes and Georg Schett. The pathogenesis of rheumatoid arthritis. *New England Journal of Medicine*, 365(23):2205–2219, 2011.

[68] A. McLean-Tooke, C. Aldridge, S. Waugh, G.P. Spickett, and L. Kay. Methotrexate, rheumatoid arthritis and infection risk — what is the evidence? *Rheumatology*, 48(8):867, 2009.

[69] D.J. Michael and A.C. Nelson. Handx: A model-based system for automatic segmentation of bones from digital hand radiographs. *Medical Imaging, IEEE Transactions on*, 8(1):64–69, 1989.

[70] R. Paul Mihail, Gustav Blomquist, and Nathan Jacobs. A CRF approach to fitting a generalized hand skeleton model. In *IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2014.

[71] R. Paul Mihail, Nathan Jacobs, and Judy Goldsmith. Real time gesture recognition with 2 kinect sensors. In *International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV)*, 2012.

[72] G. Nesher, T.G. Osborn, T.L. Moore, et al. Effect of treatment with methotrexate, hydroxychloroquine, and prednisone on lymphocyte polyamine levels in rheumatoid arthritis: correlation with the clinical response and rheumatoid factor synthesis. *Clinical and experimental rheumatology*, 15(4):343, 1997.

[73] L. Nielsen-Bohlman. *Health literacy: A prescription to end confusion*. National Academy Press, 2004.

[74] C. Nøhr and J. Aarts. Use of serious health games in health care: a review. *Information Technology in Health Care: Socio-Technical Approaches 2010: From Safe Systems to Patient Safety*, 157:160, 2010.

[75] A.M. O'Connor, C.L. Bennett, D. Stacey, M. Barry, N.F. Col, K.B. Eden, V.A. Entwistle, V. Fiset, M. Holmes-Rovner, S. Khangura, H. Llewellyn-Thomas, and Rovner D. Decision aids for people facing health treatment or screening decisions. *Cochrane Database of Systematic Reviews*, 3(CD001431), 2009.

[76] A.M. O'Connor, I.D. Graham, and A. Visser. Implementing shared decision making in diverse health care systems: the role of patient decision aids. *Medical Decision Making*, 57(3):247–249, June 2005.

[77] A.M. O'Connor, F. Legare, and D. Stacey. Risk communication in practice: the contribution of decision aids. *BMJ*, 327(6):736–740, 2003.

[78] A.M. O'Connor, D Stacey, V Entwistle, H Llewellyn-Thomas, D Rovner, M Holmes-Rovner, V Tait, J Tetroe, V Fiset, M Barry, et al. Decision aids for people facing health treatment or screening decisions (review). *The Cochrane Library*, 4, 2004.

[79] L. Osterberg and T. Blaschke. Adherence to medication. *New England Journal of Medicine*, 353(5):487–497, 2005.

[80] Mikkel Østergaard, Michael Hansen, Michael Stoltenberg, Karl Erik Jensen, Marcin Szkudlarek, Brigitta Pedersen-Zbinden, and Ib Lorenzen. New radiographic bone erosions in the wrists of patients with rheumatoid arthritis are detectable with magnetic resonance imaging a median of two years earlier. *Arthritis & Rheumatism*, 48(8):2128–2131, 2003.

[81] T. Pincus, G. Ferraccioli, T. Sokka, A. Larsen, R. Rau, I. Kushner, and F. Wolfe. Evidence from clinical trials and long-term observational studies that disease-modifying anti-rheumatic drugs slow radiographic progression in rheumatoid arthritis: updating a 1983 review. *Rheumatology*, 41(12):1346–1356, 2002.

[82] T. Pincus, A. Larsen, R.H. Brooks, J. Kaye, E.P. Nance, L.F. Callahan, et al. Comparison of 3 quantitative measures of hand radiographs in patients with rheumatoid arthritis: Steinbrocker stage, kaye modified sharp score, and larsen score. *The Journal of rheumatology*, 24(11):2106, 1997.

[83] Theodore Pincus and Tuulikki Sokka. Quantitative measures for assessing rheumatoid arthritis in clinical trials and clinical care. *Best Practice & Research Clinical Rheumatology*, 17(5):753–781, 2003.

[84] J. Protheroe, P. Bower, C. Chew-Graham, J.T. Peters, and T. Fahey. Effectiveness of a computerized decision aid in primary care on decision making and quality of life in menorrhagia: Results of the mentip randomized controlled trial. *Medical Decision Making*, 27(5):575–584, 2007.

[85] T.M. Rader, L. Ghogomu, E. Tugwell, and P. Welch. A cochrane decision aid to discuss options with your doctor. Februrary 2011.

[86] L. Reichlin, N. Mani, K. McArthur, A.M. Harris, N. Rajan, and C.C. Dacso. Assessing the acceptability and usability of an interactive serious game in aiding treatment decisions for patients with localized prostate cancer. *J Med Internet Res*, 13(1), Jan 2011.

[87] Z. Ren, J. Meng, J. Yuan, and Z. Zhang. Robust hand gesture recognition with kinect sensor. In *Proceedings of the 19th ACM international conference on Multimedia*, pages 759–760. ACM, 2011.

[88] V.F. Reyna, W.L. Nelson, Han P.K., and Dieckmann N.F. How numeracy influences risk comprehension and medical decision making. *Psychological Bulletin*, 135(6):943, 2009.

[89] K.G. Saag, G.G. Teng, N.M. Patkar, J. Anuntiyo, and C. Finney. American college of rheumatology 2008 recommendations for the use of nonbiologic and biologic disease-modifying antirheumatic drugs in rheumatoid arthritis. *Arthritis Care & Research*, (59):762–784, 2008.

[90] P. Sand, L. McMillan, and J. Popović. Continuous capture of skin deformation. *ACM Transactions on Graphics (TOG)*, 22(3):578–586, 2003.

[91] M.M. Schapira, A.B. Nattinger, and T.L. McAuliffe. The influence of graphic format on breast cancer risk communication. *Journal of Health Communication*, 11, 2006.

[92] A.P. Schwab. Putting cognitive psychology to work: Improving decision-making in the medical encounter. *Social Sciences and Medicine*, 67(11):1861–1869, December 2008.

[93] Jasvinder A. Singh, Daniel E. Furst, Aseem Bharat, Jeffrey R. Curtis, Arthur F. Kavanaugh, Joel M. Kremer, Larry W. Moreland, James O'Dell, Kevin L. Winthrop, Timothy Beukelman, et al. 2012 update of the 2008 american college of rheumatology recommendations for the use of disease-modifying antirheumatic drugs and biologic agents in the treatment of rheumatoid arthritis. *Arthritis care & research*, 64(5):625–639, 2012.

[94] T. Sokka. Radiographic scoring in rheumatoid arthritis. *Bulletin of the NYU Hospital for Joint Diseases*, 66(2):166–8, 2008.

[95] J.M. Sotoca, J.M. Iñesta, and M.A. Belmonte. Hand bone segmentation in radioabsorptiometry images for computerised bone mass assessment. *Computerized Medical Imaging and Graphics*, 27(6):459–467, 2003.

[96] S. Sueda, A. Kaufman, and D.K. Pai. Musculotendon simulation for hand animation. In *ACM Transactions on Graphics (TOG)*, volume 27, page 83. ACM, 2008.

[97] Poonam Suryanarayan, Anbumani Subramanian, and Dinesh Mandalapu. Dynamic hand pose recognition using depth data. In *ICPR*, pages 3105–3108, 2010.

[98] Tarja Susi, Mikael Johannesson, and Per Backlund. Serious games : An overview. Technical Report HS- IKI -TR-07-001, University of Skvde, School of Humanities and Informatics, 2007.

[99] C.Y. Tang, G. Zhang, and C.P. Tsui. A 3d skeletal muscle model coupled with active contraction of muscle fibres and hyperelastic behaviour. *Journal of biomechanics*, 42(7):865, 2009.

[100] D. Thalmann, J. Shen, and E. Chauvineau. Fast realistic human body deformations for animation and vr applications. In *Computer Graphics International, 1996. Proceedings*, pages 166–174. IEEE, 1996.

[101] Michael E. Tipping and Christopher M. Bishop. Probabilistic principal component analysis. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 61(3):611–622, 1999.

[102] L.J. Trevena, A. Barratt, O. Butow, and P. Caldwell. A systematic review on communicating with patients about evidence. *Journal of Evaluation in Clinical Practice*, 12(1):12–23, 2006.

[103] P. Tugwell, B. Shea, M. Boers, P. Brooks, L. Simon, V. Strand, and G. Wells. *Evidence-based rheumatology*, chapter 4, pages 41–56. BMJ Books, 2005.

[104] Martijn Van De Giessen, Geert J. Streekstra, Simon D. Strackee, Mario Maas, Kees A. Grimbergen, Lucas J. Van Vliet, and Frans M. Vos. Constrained registration of the wrist joint. *Medical Imaging, IEEE Transactions on*, 28(12):1861–1869, 2009.

[105] J.E. Wennberg. Unwarranted variations in healthcare delivery: implications for academic medical centres. *BMJ*, 325(7370):961, 2002.

[106] J. Wilhelms and A. Van Gelder. Anatomically based modeling. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, pages 173–180. ACM Press/Addison-Wesley Publishing Co., 1997.

[107] Jacob O. Wobbrock, D. Wilson, Andrew, and Yang Li. Gestures without libraries, toolkits or training: a $1 recognizer for user interface prototypes. In *Proceedings of the 20th annual ACM symposium on User interface software and technology*, pages 159–168. ACM, 2007.

[108] Ying Wu, Thomass S. Huang, and Thomas S. Huang. View-independent recognition of hand postures. In *In CVPR*, pages 88–94, 2000.

[109] Ying Wu, John Y. Lin, and Thomas S. Huang. Capturing natural hand articulation. In *ICCV*, pages 426–432, 2001.

[110] Eric P. Xing, Andrew Y. Ng, Michael I. Jordan, and Stuart Russell. Distance metric learning with application to clustering with side-information. *Advances in neural information processing systems*, pages 521–528, 2003.

[111] Chenyang Xu and Jerry L. Prince. Generalized gradient vector flow external forces for active contours. *Signal processing*, 71(2):131–139, 1998.

[112] Ayhan Yuksel and Tamer Olmez. Automatic segmentation of bone tissue in x-ray hand images. In *Adaptive and Natural Computing Algorithms*, pages 590–599. Springer, 2009.

[113] Xiaojin Zhu and Andrew B Goldberg. Introduction to semi-supervised learning. *Synthesis lectures on artificial intelligence and machine learning*, 3(1):1–130, 2009.

[114] B.J. Zikmund-Fisher, A. Fagerlin, T.R. Roberts, H.A. Derry, and P.A. Ubel. Patient education in rheumatoid arthritis: the effectiveness of the ARC booklet and the mind map. *Journal of Rheumatology*, (46):1593–1596, 2007.

[115] B.J. Zikmund-Fisher, A. Fagerlin, T.R. Roberts, H.A. Derry, and P.A. Ubel. Alternate methods of framing information about medication side effects: Incremental risk versus total risk of occurrence. *Journal of health communication*, 13(2):107–124, 2008.

Vita

Name: Radu P. Mihail

Education

- B.S., Computer Science, *Cum Laude, Honors Program* December, 2009
  Minors: Mathematics, Statistics

Employment

2010-2014 , TA, Fellow and RA, University of Kentucky, Lexington, Kentucky, USA

2009 Consolutant, Claraview, a Division of Teradata, Reston, Virginia, USA

2006-2009 DBA, Ashland Inc., Lexington, Kentucky, USA

2005-2006 Web developer, Eastern Kentucky University, Richmond, Kentucky, USA

Honors and Awards

- Provost Award for Outstanding Teaching Assistant, 2013
- SACM Award for Outstanding Teaching Assistant, 2013
- Halcomb Fellowship in Medicine and Engineering, 2011
- Outstanding Co-op, Ashland Inc., 2008
- President's List, Eastern Kentucky University, several semesters starting May 2007
- Presidential Scholarship, Eastern Kentucky University, January 2006
- First prize in a national programming contest, Focsani, Romania, January 2003

Publications

1. R. Paul Mihail, Gustav Blomquist, Nathan Jacobs. "A CRF Approach to Fitting a Generalized Hand Skeleton Model", Winter Conference on Applications of Computer Vision (WACV), 2014.
2. R. Paul Mihail, Beth Rubin, Judy Goldsmith. "Online Discussions: Improving Education in CS?", Proceedings of the 44th ACM technical symposium on Computer science education (SIGCSE), 2014.
3. R. Paul Mihail, Judy Goldsmith, Nathan Jacobs, Jerzy Jaromczyk, "Teaching Graphics for Games using Microsoft XNA", In International Conference on Computer Games (CGAMES), 2013.
4. Kaitlin Burton, Frederick Hallock, R. Paul Mihail, "A Data-Driven Approach to Visualize the Effects of Rheumatoid Arthritis on Hands", In International Conference on Computer Games (CGAMES), 2013.
5. R. Paul Mihail, Nathan Jacobs, Judy Goldsmith. "Real Time Gesture Recognition-With 2 Kinect Sensors", 16th International Conference on Image Processing, Computer Vision, & Pattern Recognition (IPCV), 2012.

Presentations

- "Teaching Graphics for Games using Microsoft XNA", Presented at CGames 2013, Louisville, KY, July 2013
- "Static Hand Gesture Recognition with 2 Kinect Sensors", Presented at Worldcomp 2012, Las Vegas, NV, July 2012
- "Decision making among patients with lower literacy and numeracy", Presented at the AI seminar at University of Kentucky, June 2010
- "Machine Learning and Artificial Intelligence: An Overview", Presented at the Kentucky Honors Roundtable in February 2009 at Murray State University, September 2009

Service

- Student Member of the Higher Degrees Committee for the Computer Science Department, 2013
- Student Member of the External Review Committee for the Computer Science Department, 2012
- University of Kentucky Graduate School Microteaching Leader, August 2011
- Student Member of the Search Committee for the Dean of the CoE, November 2011