Applying Video Magnification Techniques to the Visualization of Blood Flow
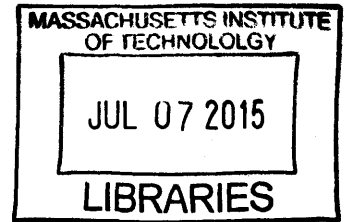
by

Amy (Xiaoyu) Zhao

Submitted to the Department of Electrical Engineering and Computer
Science
in partial fulfillment of the requirements for the degree of

Master of Science in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2015

© Massachusetts Institute of Technology 2015. All rights reserved.

Author . . . .
Signature redacted

Department of Electrical Engineering and Computer Science
May 20, 2015

Certified by . .
Signature redacted

John V. Guttag
Dugald C. Jackson Professor
Thesis Supervisor

Certified by . . . .
Signature redacted

Frédo Durand
Professor
Thesis Supervisor

Accepted by . . .
Signature redacted

/ Professor Leslie A. Kolodziejski
Chair of the Committee on Graduate Graduate Students

Applying Video Magnification Techniques to the Visualization of Blood Flow

by

Amy (Xiaoyu) Zhao

Abstract

In this thesis, we investigate the use of video magnification for the visualization and assessment of blood flow. We address the challenge of low signal-to-noise ratios in video magnification by modeling the problem and developing an algorithm for measuring the SNR in the context of video magnification. We demonstrate that the algorithm can be used to estimate the SNR of a real video and predict the SNR in the magnified video. We use several techniques based on video magnification to visualize the blood flow in a healthy hand and a hand with an occluded artery, and show that these visualizations highlight differences between the hands that might be indicative of important physiological differences.

Thesis Supervisor: John V. Guttag
Title: Dugald C. Jackson Professor

Thesis Supervisor: Frédo Durand
Title: Professor

Acknowledgments

I would like to express my sincere gratitude to the individuals who made this work possible:

Prof. John V. Guttag, my research advisor and thesis supervisor, for spending many hours in the proofreading of this document, but also for constantly encouraging me to improve in all aspects of my work.

Prof. Frédo Durand, my thesis supervisor, for his insightful technical analyses of my work, but also for pushing me to challenge myself and to "do good research" above all else.

Prof. Collin M. Stultz, for his considerable medical expertise, but also for teaching me to question everything.

I would like to thank the students of the Data Driven Inference Group and the students of the Computation Biophysics Group for generously lending me their feedback and encouragement over the years.

As always, I am grateful to my mother, Rachel; my father, Qihua; and my sister, Linda for their unwavering love and support.

I would like to thank the members of my household: Robert, Mocha, and Mochi, for giving me the best home that I could ask for.

I am grateful to Dr. Irwin Jacobs, Mrs. Joan Jacobs, and Quanta Computer for their support.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Video magnification technology allows us to visualize small variations in videos that are difficult or impossible to see with the naked eye. These techniques can be used to investigate a wide range of physical phenomena. For instance, Eulerian linear video magnification has been used to visualize the subtle motion in the chest of a breathing baby [1]. Recently, new phase-based algorithms have been used to visualize the vibration of mechanical structures [2, 3].

In addition to magnifying motion, the Eulerian linear video magnification algorithm magnifies color changes. It has been used to produce visualizations of the subtle green to red variations that occur at the heart rate as blood perfuses the human face [1]. Blood flow measures play a large role in the evaluation of peripheral vascular diseases [4]. This thesis investigates the use of video magnification for the visualization and assessment of blood flow, and its potential applications in evaluating vascular disease. In this work, we examine the accuracy of the visualizations produced by video magnification. We demonstrate that video magnification techniques can produce visualizations that exhibit noticeable differences between a healthy subject and a subject affected by an arterial occlusion.

In this chapter, we discuss the clinical motivation for visualizing blood flow. We also explore previous work in obtaining measures of blood flow.

1.1 Clinical motivation

In 2010, approximately 8 million adults in the US and 202 million people globally were affected by peripheral arterial disease (PAD) [5, 6]. PAD is a cardiovascular condition that is characterized by insufficient blood flow in the organs and limbs resulting from a narrowing of the peripheral arteries. PAD can affect any of the arteries outside of the coronary arteries, including vessels in the brain, upper extremities and kidneys [7], but most commonly affects the lower extremities [8]. While the majority of PAD cases are asymptomatic, asymptomatic and symptomatic PAD are both associated with increased risk of adverse vascular events and increased mortality [9, 10]. Patients with symptomatic PAD experience decreased quality of life due to symptoms such as claudication (reproducible muscle pain related to physical activity), pain at rest, non-healing wounds and ulcerations. Treatments for severe symptoms of PAD such as critical limb ischemia are typically invasive, such as revascularization or amputation [11]. PAD is a growing problem in the US and other countries with aging populations, since the prevalence of PAD is age-related [12]. Other major risk factors for PAD include smoking, diabetes and hypertension [4].

1.1.1 Arterial blood flow

Before we delve any further into PAD, let us briefly review human vascular anatomy. In the human systemic circulation system, arteries are the blood vessels that carry blood away from the heart. These vessels branch into a network of successively smaller vessels. As an artery enters the organ that it is supplying, it branches into arterioles, which in turn branch into a network of capillaries, the smallest class of vessels. These capillaries rejoin to become venules and then veins that carry blood away from the organ and back into the heart [13].

The heart contracts rhythmically in a pattern called the cardiac cycle, which consists of two main phases: systole and diastole. During systole, the heart contracts; during diastole, the heart refills with blood. This process causes blood to be ejected in a pulsatile manner. The arterial wall is elastic and responds to propagated pulse
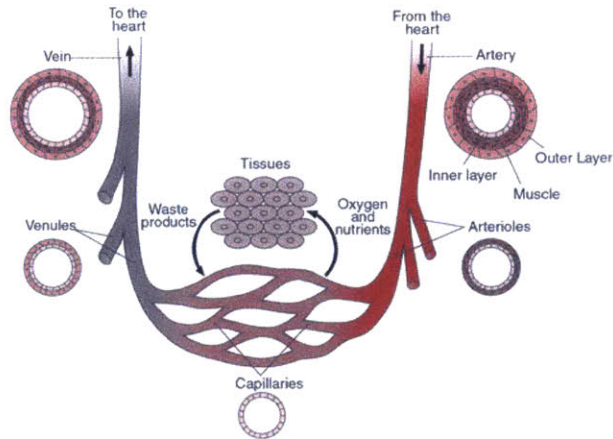
Figure 1-1: Arteries carry blood from the heart to the tissues of the body. As arteries near the organs, they branch into arterioles and then capillaries, through which the exchange of oxygen and nutrients occur. Venules and veins carry the deoxygenated blood back to the heart. Figure reproduced from [14].

of blood by increasing in diameter [15]. This dampens the pulsatile energy of the blood as it approaches the organs. In the arterioles and capillaries, the flow of blood is largely non-pulsatile [16].
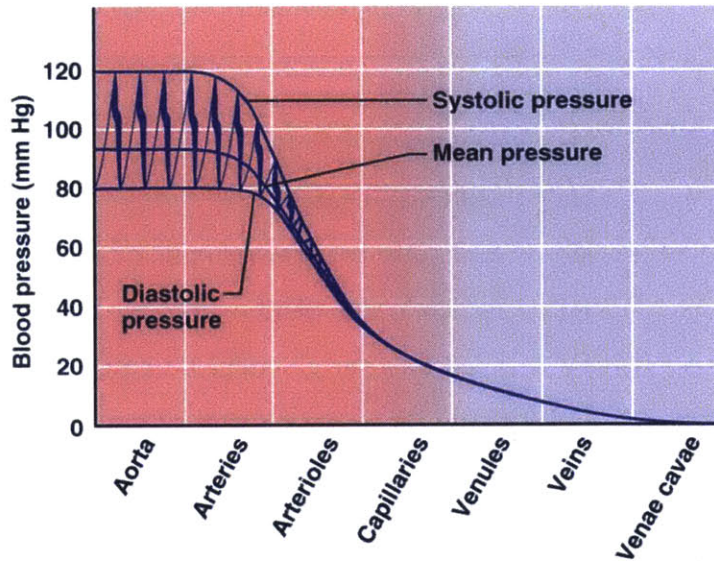
Figure 1-2: The aorta and the arteries of the body experience different blood pressures during systole and diastole. This pulsatile flow is dampened by the arterial walls; blood flow is mostly non-pulsatile by the time it enters the smaller vessels. Figure reproduced from [17].

The pulsatile quality of arterial flow serves as the basis for many measures of

17

vascular health. For instance, the shape of the blood velocity waveform measured by Doppler ultrasonography (which we will discuss in section 1.1.3) can indicate the presence of an arterial obstruction.

1.1.2   Causes and symptoms of PAD

PAD is most commonly caused by atherosclerosis [4], a pathologic process in which the arteries become narrower due to the accumulation of lipid and fibrous materials in the vessel walls [18]. Other causes of PAD include arterial aneurysms, thrombosis, arterial injury, etc. [19].

The severity of the symptoms experienced by an individual with PAD depends on the degree of stenosis (that is, arterial narrowing), the number of affected arteries, and the patient's physical activity level. Many patients with PAD do not experience any symptoms. Patients with more advanced PAD experience symptoms ranging from intermittent claudication (pain during exercise) to non-healing ulcers and gangrene [19].

Stage	Symptom
I	Asymptomatic
IIa	Mild claudication
IIb	Moderate to severe claudication
III	Ischemic rest pain
IV	Ulceration or gangrene

Table 1.1: Fontaine's Stages for the classification of PAD. Patients are classified as asymptomatic if they do not exhibit typical claudication symptoms. Claudication refers to the reproducible muscle discomfort that produced by exercise and relieved by resting for 10 minutes [20]. Ischemic rest pain refers to pain at rest that is caused by insufficient blood supply [19]. While this classification method is not used in clinical practice, it is commonly used for research purposes. Table reproduced from [20].

It is estimated there there are 2 to 5 times as many asymptomatic individuals with PAD as symptomatic individuals. Individuals are designated as asymptomatic

if they have no symptoms, or if they do not exhibit any typical claudication symptoms. However, asymptomatic individuals may have more subtle impairments of leg function [4]. Furthermore, asymptomatic PAD and symptomatic PAD alike are associated with increased risk of mortality and cardiovascular events [9]. There has been increasing interest in detecting PAD in asymptomatic patients through routine screening procedures. We will discuss this further in the next section, where we review existing tools for the assessment and treatment of PAD.

1.1.3 Diagnostic tools for PAD

There are many tools that are currently used for diagnosing PAD. These tools range from inexpensive and easy-to-obtain blood pressure ratios, to blood flow measures that require specialized equipment in a vascular laboratory, to vascular imaging techniques that require intravenously injected contrast agents. Video magnification holds the potential to be less expensive and more informative than many of existing physiologic tests. In this section, we review the existing techniques and their associated challenges.

Non-invasive vascular measures

There are several non-invasive examinations that can be used to diagnose PAD, and to determine the location and severity of stenoses. The ankle-brachial index (ABI) is a common test that is inexpensive and easy to perform. The ABI serves as the standard for diagnosing lower extremity PAD in office practice [4]. It is also recommended as a screening device for asymptomatic patients who are at risk for PAD [20, 21]. It is computed by measuring the systolic blood pressure from the brachial artery (in the upper arm) of both arms, and the dorsalis pedis (DO) and posterior tibial (PT) arteries (in the lower calf). The ABI is defined as the ratio of the higher of the two pressures from the ankle to the higher of the pressures from the two arms. Similar measures include the wrist-brachial index and toe-brachial index, which are computed in a similar fashion to the ABI. The wrist-brachial index is used to diagnose upper

19

extremity PAD, and the toe-brachial index is more useful than the ABI in certain patient populations with lower extremity PAD [22].
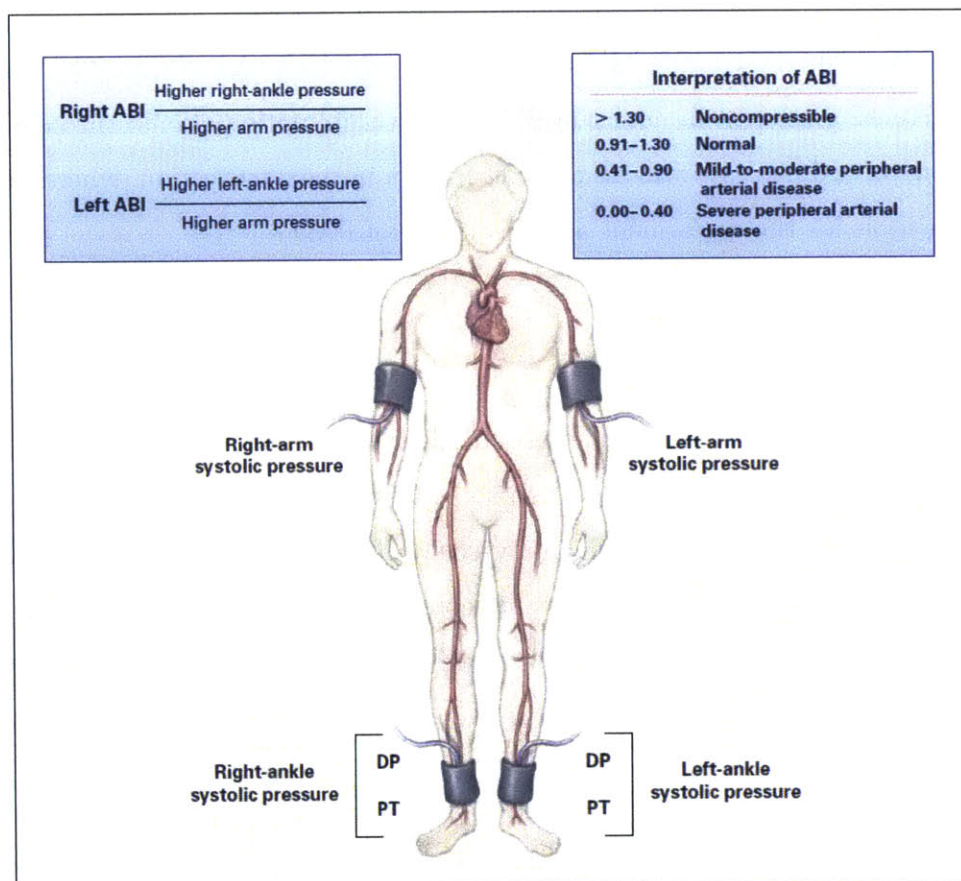
Figure 1-3: The ABI is computed by measuring the systolic blood pressure the dorsalis pedis (DP) and posterial tibila (PT) arteries in each ankle, and the pressure in each arm. These measurements are usually made using blood pressure cuffs and a Doppler ultrasound device. The ABI is computed by dividing the higher ankle pressure in each leg by the higher arm pressure. An ABI > 1.3 suggests that vessels are calcified, and that additional vascular studies such as toe-brachial indices or pulse volume recordings should be conducted [22]. Figure reproduced from [23].

A diagnosis made by the ABI is often considered accurate enough that verification with further diagnostic techniques is not required. However, the ABI produces inaccurate results in some patient populations; since the process of measuring blood pressure requires using a blood pressure cuff to abolish systolic blood flow, the ABI is falsely elevated when the patient has calcified vessels that do not compress normally. This occurs in a small fraction of diabetic or elderly individuals. In these popula-

tions, other diagnostic measures such as the toe-brachial index (which is useful since the small vessels in the toes are frequently spared from calcification) or pulse volume recordings (which we will discuss below) may provide more reliable information [22].

Once a diagnosis of PAD is made using the ABI, further studies are often required to determine the location and severity of the disease. Segmental pressure examinations and pulse volume recordings (PVR) use specialized equipment in a vascular laboratory to measure blood pressure and limb volume respectively at successive locations in an extremity [22]. Continuous-wave Doppler ultrasound can similarly be used to identify the location, severity and progression of PAD. A continuous-wave Doppler ultrasound device works by continuously emitting sound waves and measuring the reflected waves in order to compute the real-time blood flow velocity [24, 21]. This device is commonly used to aid in the measurement of the ABI. Qualitative and quantitative measures of the continuous-wave Doppler ultrasound waveform also provide information about the severity of PAD [21]. Similar to the ABI, segmental pressure examinations and continuous-wave Doppler ultrasound measurements have limited accuracy in certain patient populations, including patients with highly calcified and incompressible arteries [4].

Vascular imaging

Vascular imaging studies are sometimes needed to further assess PAD anatomy, particularly in patients who are candidates for invasive treatments [4]. Cather angiography, or contrast angiography (CA) is the gold standard for vascular imaging. It is capable of producing high-resolution images of vascular anatomy. It requires the use of an injected contrast agent in conjunction with imaging technology. The contrast agent is administered via a percutaneous catheter inserted near the target vessel to be imaged. CA is an invasive procedure and incurs risks associated with percutaneous arterial access, catheter manipulation, contrast agents and radiation exposure. It is also the most costly imaging technique in use for PAD [25]. There are several non-invasive alternatives to CA. Duplex ultrasound is a relatively low-cost and widely used imaging technique [26]. It combines continuous-wave Doppler ultrasound with B-mode

ultrasound, which uses reflected sound waves to produce grayscale images of tissue structures. Duplex ultrasound can be used to visualize the flow of blood in vascular structures [22]. Magnetic resonance angiography (MRA) is another non-invasive imaging technique, and studies suggest that it is more accurate than duplex ultrasound [26]. MRA uses magnetic resonance imaging technology, often coupled with intravenously injected contrast-enhancing agents, to rapidly produce two-dimensional and three-dimensional images of arterial structures. Computed tomographic angiography (CTA) is a relatively new non-invasive imaging technique, and has been shown in several early studies to perform well compared to existing techniques for detecting occlusions [26]. CTA uses an intravenously injected contrast agent in conjunction with X-ray imaging. Like MRA, CTA is capable of rapidly acquiring images to produce three-dimensional images of arterial structures.

	Diagnostic tool	Provides information about	Advantages and disadvantages
Non-invasive vascular measures	Ankle-brachial index (ABI)	Severity	✓ Fast, easy to perform ✗ May be inaccurate for non-compressible arteries
	Segmental pressure examination	Severity, location	✗ May be inaccurate for non-compressible arteries
	Pulse volume recording (PVR)	Severity, location	✓ Useful for non-compressible arteries ✗ Less accurate than other noninvasive localization tests
	Continuous-wave Doppler ultrasound	Severity, location	✓ Useful for non-compressible arteries ✗ May report normal waveform downstream of severe stenoses ✗ Limited accuracy in overlapping or densely calcified segments
Vascular imaging techniques	Duplex ultrasound	2D vascular structure	✓ Inexpensive compared to CA ✓ Low-risk
	Magnetic resonance angiography (MRA)	3D vascular structure	✓ Fast ✓ Inexpensive compared to CA ✗ Cannot be used in patients with stents, pacemakers, and certain implants
	Computed tomography angiography (CTA)	3D vascular structure in volumetric slices	✓ Fast ✓ Inexpensive compared to CA ✗ Risks associated with contrast agents ✗ Relatively new, requires more studies of accuracy and effectiveness
	Contrast angiography (CA)	2D vascular structure in limited orientations	✓ Established technique with well-studied accuracy and effectiveness ✗ Expensive ✗ Invasive ✗ Risks associated with contrast agents

Table 1.2: Comparison of diagnosis and imaging tools for PAD. These tools are listed in order from least expensive (ABI) to most expensive (CA). Many of these tools are also useful for monitoring patients after therapeutic interventions such as revascularization. Information adapted from [4, 22, 20].

The diagnostic tools that are currently used for PAD are summarized in Table 1.2. We hypothesize that video magnification technologies have the potential to visualize blood flow in a way that can identify the severity and location of PAD. These visualizations are likely to be easier and less expensive to obtain than segemental pressure examinations, PVRs, and even the ABI. Furthermore, video magnification may be able to visualize the flow in vessels that are too small for existing vascular imaging techniques to capture. We will explain this in the following section, where we review some existing video-based techniques for assessing blood flow, and in particular, some existing video magnification algorithms.

1.2 Video-based techniques for measuring blood flow

The idea of obtaining measures of blood flow from videos is not new. In the past several decades, there has been considerable interest in simple and low-cost methods for obtaining measures of cardiac health; video-based methods have been particularly sought-after as they are inexpensive and non-invasive. Recently, Verkruysse *et al.* demonstrated that is possible to obtain measurements of blood flow from videos taken with a consumer grade digital camera [27]. Their method is based on photoplethysmography (PPG), a simple and inexpensive optical technique in which the skin is illuminated with light of specific wavelengths (typically red or infrared), and a photodetector is used to measure changes in the reflected light that are associated with changes in blood volume. PPG utilizes the principle that light is attenuated by its interactions with biological tissue; the amount of light received by the photodetector is dependent on factors such as the blood volume in the tissue, the movement of the blood vessel walls, and the oxygenation of blood [28]. In their work, Verkruysse *et al.* took a video of a person's face under ambient lighting. From the video, they measured the average intensity of a user-defined region of interest containing the face. After applying spatial averaging and temporal filtering to improve the signal-to-noise ratio, they were able to obtain accurate measures of the heart rate, as well as visualizations of the pulse amplitude and phase of blood flow in the face. In [29], Poh *et al.*

obtained measures of the average heart rate from videos of stationary human faces. Their approach examined the mean intensity of an automatically determined region of interest in a video of a person's face, and applied blind source separation algorithms to obtain the signal corresponding to the flow of blood in the skin. Balakrishnan *et al.* showed that it is also possible to obtain a signal related to blood flow by examining the motion of the head [30]. Their approach works by examining the horizontal and vertical trajectories of multiple tracked points in the face. Blind source separation techniques are then applied to produce a signal that coincides withe the flow of blood in the head, as verified by an electrocardiogram.

1.3   Video magnification for visualizing blood flow

We can view video magnification algorithms as a special class of video-based techniques. In general, the goal of video magnification is to magnify small changes in videos that are not visible to the naked eye. When applied to the task of examining blood flow, video magnification produces visualizations rather than measurements of the variations corresponding to the flow of blood.

Early work in video magnification took a Lagrangian approach, in which motions are explicitly computed and video frames are warped according to the magnified motion vectors [31, 32]. These techniques are useful for magnifying small motions, but are computationally expensive and often produced noticeable artifacts [2]. In recent years, several Eulerian approaches were proposed, in which the variations at specific pixel locations are amplified [1, 2, 3]. These approaches are computationally inexpensive, and have been shown to be particularly adept at visualizing repetitive motions. The linear Eulerian video magnification algorithm (which we will refer to as linear video magnification) amplifies the intensity changes of pixels over time. This algorithm is useful for magnifying color changes as well as motion; for instance, it has been used to visualize the flow of blood in a human face [1]. However, the algorithm also magnifies some of the noise in the input video [2]. Eulerian phase-based motion magnification algorithms amplify the changes in spatial phase of pixels over time

[2, 3]. The algorithms are capable of producing videos with convincing and relatively noise-free magnified motions. For instance, they have been used to visualize the movement of a digital camera during the closing of the shutter [2], and the vibration of mechanical structures after being subjected to an impact [33]. In this work, we focus primarily on linear video magnification and its ability to visualize the color changes that occur as a result of blood perfusion in tissues.

The linear video magnification algorithm is capable of amplifying color changes and motion in videos taken with consumer grade cameras, using software running on a consumer grade computer. From an input video, it isolates the changes of interest by applying a spatial low-pass filter and a temporal bandpass filter. The application of the spatial lowpass filter is equivalent to performing spatial averaging on each frame, and serves the purpose of reducing the noise level and improving the signal-to-noise ratio (we will discuss this further in the next chapter). The temporal bandpass filter is used to isolate the frequencies close to the human heart rate. These changes are then multiplied by some constant factor and recombined with the input video to create an output video in which the isolated color changes and motions are amplified. Using this algorithm, Wu *et al.* were able to visualize the repetitive green to red variations that occur when blood perfuses the face [34].

These results indicate that Eulerian linear video magnification may be useful for assessing blood perfusion in tissues. We believe that Eulerian linear video magnification can contribute to the diagnosis of PAD in two ways:

Cost   Video magnification can quickly and cheaply produce visualizations that can aid in identifying the location and severity of PAD. Such a tool may be used as a low-cost alternative for the existing non-invasive vascular measures used in diagnosis of PAD, such as the ABI or PVRs.

Aggregate flow information

We believe that Eulerian linear video magnification can provide information about the blood flow patterns in blood vessels that are too small to be captured by existing imaging techniques. Eulerian linear video magni-

fication produces visualizations of the color changes that occur near the surface of the skin; these color changes reflect the blood volume changes in the small blood vessels[1] close to the surface [27, 28]. Existing vascular imaging techniques are adept at visualizing the structure of large blood vessels, but are not suitable for imaging small vessels. For instance, high-resolution MRA is estimated to have a spatial resolution of $0.8mm \times 0.8mm$ [37]; for comparison, the average diameter of capillaries in the human hand is estimated to be $10\mu m$ [38]. The blood flow visualizations produced by video magnification might be used to supplement the information obtained by vascular imaging techniques.

While Eulerian linear video magnification holds promise, at present, it also faces some challenges. Eulerian video magnification sometimes produces videos in which the signal-to-noise ratio (SNR) that is too low for the visualizations to be accurate or reliable. A low output SNR is caused by two main factors. The first factor is low signal amplitudes in the input video. Video magnification techniques produce the most interesting results when applied to small variations that are not discernible by the naked eye. Thus, the signal of interest almost always has a small amplitude compared to the image content. The second factor is the presence of noise that is difficult to remove. Photoplethysmographic measurements are susceptible to motion-related signal corruption [39]; these unwanted variations often fall within the same temporal frequency band as the physiological signal of interest [29]. For instance, we may be interested in visualizing the rate at which blood perfuses the head, but as shown in [30], the flow of blood in the head is associated with a "bobbing" motion that is likely to obscure the more subtle variations caused by perfusion. Such unwanted variations cannot be removed using temporal filtering alone. Another noise term that is difficult to remove is the acquisition noise that is introduced by the video capture device, typically a digital camera. As we will discuss later, noise in digital cameras is generally assumed to be white Gaussian noise. This can be attenuated through spatial

---

[1]While the flow of blood in small blood vessels (or, the microvasculature) is generally assumed to be steady, studies have shown that the flow in some human vessels of inner diameter $\sim 20\mu m$ still have pulsatile qualities [35, 36].

and temporal filtering but is difficult to remove completely. In video magnification algorithms, noise that is not removed by the filtering steps will be magnified in the output video. The combination of noise in the output and low signal amplitudes may result in an SNR that is too low for video magnification to produce meaningful results. To our knowledge, no studies thus far have attempted to measure the SNR in the context of video magnification.

In the following chapters, we present a a framework for determining the SNR of an input video that is intended for video magnification. We discuss a model of the signal and noise suitable for video magnification in general, as well a model for the specific use of video magnification for visualizing blood flow. We also present two algorithms for measuring the SNR of an input video, and compare the algorithms using a simulated image sequence.

Chapter 2

Signal and noise in video magnification

In applications of video magnification, the visual changes of interest are often small relative to the noise level. This low signal-to-noise ratio (SNR) raises the following question: how accurate are the visualizations produced by video magnification algorithms? In particular, what is the SNR in visualizations of blood flow? In this chapter, we address this question by investigating the following sub-problems:

1. Model the signal and noise in a video intended for video magnification. This model will allow us to develop algorithms for the next two steps.

2. Measure the signal amplitude and noise level in the input video. Compute the SNR of the input video (which we refer to as the input SNR).

3. Given the input SNR, predict the output SNR that can be achieved by the Eulerian linear video magnification algorithm.

To our knowledge, no studies thus far have attempted to measure the SNR of the videos used in video magnification. While this work is tailored to the application of visualizing blood flow, the framework and algorithms that we present are useful in other applications of video magnification as well.

2.1 Definitions

Video magnification is useful for visualizing small temporal variations in videos that would otherwise not be discernible. Variations in pixel intensity can originate from processes in the acquisition device such as analog-to-digital conversion (which introduces quantization noise to each pixel), or physical variations in the video scene such as color changes, lighting intensity changes or motion. Usually, one is interested in only magnifying changes in the video scene that are produced by some physical phenomenon of interest. Examples of interesting phenomena include subjecting a metal beam to an impact, or the flow of blood in the face.
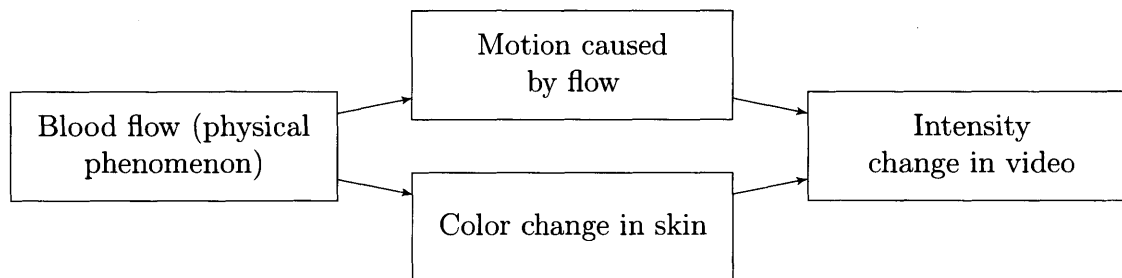
Figure 2-1: Video magnification algorithms visualize the phenomenon of interest by magnifying specific changes in the captured video. Blood flow is an example of a phenomenon that causes two types of changes in the video scene. In our work, we only want to visualize the color change in the skin.

We use the term **source of interest** to refer to the physical phenomenon producing the visual changes that one wishes to visualize. As described in Fig. 2-1, this phenomenon can produce multiple types of changes in the video scene. We define the intensity variations caused by the changes that we wish to visualize as the **signal of interest** (which we shall henceforth refer to as just the **signal**), and intensity variations caused by the the changes that we do not wish to visualize as **ancillary signal**. In addition to the source of interest, the scene might contain other sources of visual changes. These visual changes are also captured by the acquisition device as intensity changes in the video. The intensity variations in a video fall into three categories: noise caused by the camera's image acquisition process (**acquisition noise**), variations caused by the phenomenon of interest, and variations caused by other

phenomena (**artifactual noise**). These terms are summarized in Fig. 2-2 below.

```
                    ┌─────────────────────┐
                    │  Temporal variations│
                    │  in acquired video  │
                    └─────────────────────┘
          ┌──────────────────┼──────────────────┐
┌──────────────────────┐ ┌──────────────────┐ ┌──────────────────────┐
│ Noise caused by camera│ │ Visual changes in│ │  Variations caused   │
│ (acquisition noise)   │ │  scene caused by │ │   by other sources   │
│                       │ │ source of interest│ │ (artifactual noise)  │
└──────────────────────┘ └──────────────────┘ └──────────────────────┘
              ┌──────────────────┴──────────────────┐
┌──────────────────────┐              ┌──────────────────────┐
│ "Interesting" vari-  │              │ Other variations caused│
│ ations caused by     │              │ by source of interest │
│ source of interest   │              │  (ancillary signal)   │
│ (signal of interest) │              └──────────────────────┘
└──────────────────────┘
```
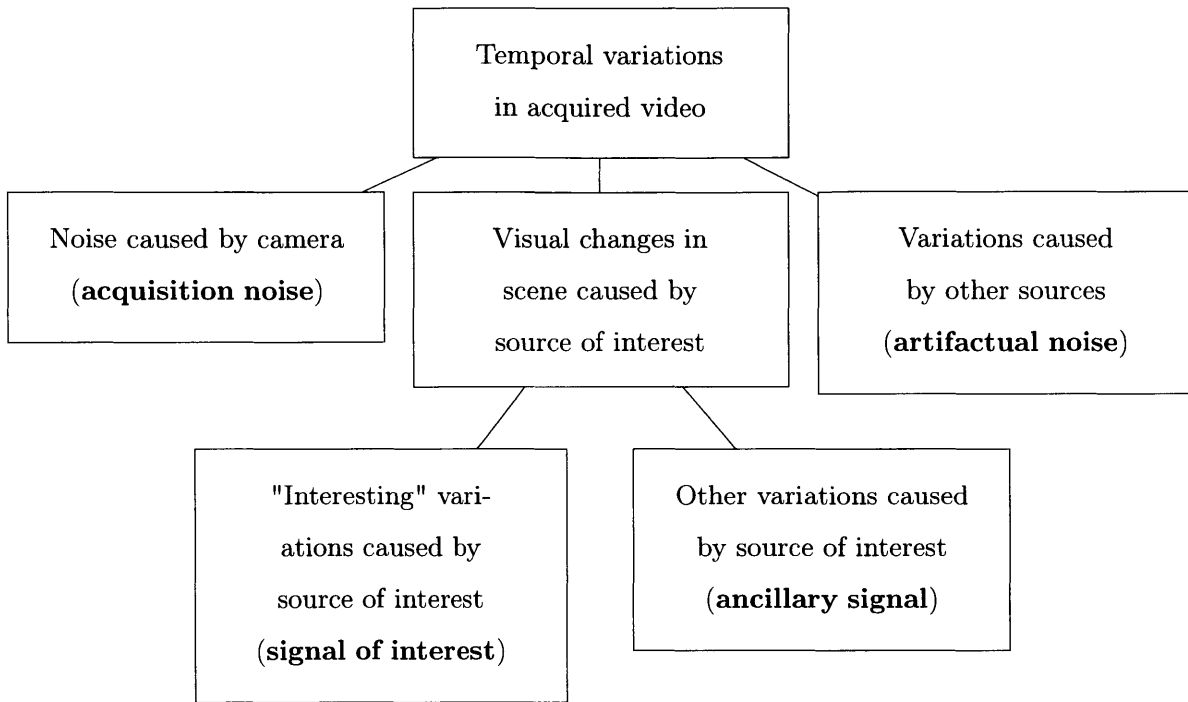
Figure 2-2: The types of temporal variations that occur in a video.

Acquisition noise

All digital cameras introduce noise during the image acquisition process. The characteristics of this type of noise have been well studied, and there are many methods for removing or attenuating it.

Signal of interest

Video magnification differs from traditional image and video processing in that the signal of interest is not the uncorrupted image or image sequence, but rather the temporal variations in the image sequence that are caused by some physical phenomenon. This signal has a small amplitude relative to the possible range of pixel intensities, and relative to the total noise level. This makes the input SNR lower than what is typically seen in the image processing literature. In some applications, such as the visualization of blood flow, the signal of interest occupies a narrow range of temporal frequencies.

31

Ancillary signal

> In some cases, the source of interest introduces additional variations that we do not wish to magnify. For example, when examining blood flow in the human hand, we might be interested in the color changes resulting from the amount of blood in the skin, but not in the motions that are caused by the mechanical effects of the blood flow. This signal can be difficult to separate from the signal of interest because it has similar temporal signatures.

Artifactual noise

> Other sources in the video scene may introduce variations to the video. For example, when examining blood flow in the human head, we may be interested in the motion caused by the flow of blood into the head, but not in the blinking motion of the eyes.

The goal of video magnification is to magnify as much of the signal and as little of the noise as possible. In other words, video magnification aims to attain the highest possible signal-to-noise ratio (SNR) in the magnified video. There are two main challenges in achieving this goal:

1. Ancillary signal and artifactual noise can be difficult to model, and cannot be removed or attenuated by existing video denoising algorithms that focus on removing acquisition noise.

2. The signal amplitude is often significantly lower than the acquisition noise level. Existing video magnification algorithms use temporal and spatial filtering to reduce the total noise level prior to magnifying the video. When the input SNR is very low, the existing denoising processes used in video magnification may be insufficient for generating an acceptable output SNR.

Later in this chapter, we shall see that in the context of visualizing blood flow, many types of ancillary signal and artifactual noise can be minimized by making fairly simple adjustments in the video capture stage or in the data processing stage. In the

next section, we discuss the problem of acquisition noise and how to measure it for the purpose of computing the input SNR.

2.2 Related work

2.2.1 Noise in images and image sequences

Modern images and image sequences are typically impaired by noise introduced by the acquisition process, the compression process, or other factors such as motion [40]. Digital cameras introduce noise at several stages in the image acquisition process. All modern digital cameras use an imaging sensor, most commonly a charge-coupled device (CCD) sensor or a complementary-metal-oxide semiconductor (CMOS) sensor to produce voltages based on the photons incident upon the sensor. The number of photons at a collection site fluctuates randomly as a result of the quantum nature of light; this introduces shot noise, which is proportional to the square root of the number of incoming photons [41]. The sensor itself introduces several types of noise including dark current noise (caused by the free electrons generated by thermal energy in the sensor [41]), readout noise (introduced by the on-chip circuit that transforms the charge at each site on the sensor to a measurable voltage [41]) and pattern noise (a systemic distortion caused by imperfections in the sensor [42]). An analog-to-digital converter is used to sample the voltages from the sensor into digital signals, adding quantization noise [43]. The intensity of a pixel in this raw image can be described as follows (see Table 2.1 for definitions):

$$I_{raw} = L + n_{shot}(\sqrt{L}) + L \cdot n_{PRNU} + n_{dark} + n_{read} + n_{FP} + n_q$$

where L is the incident light power [43, 45].

Many digital cameras apply post-capture processing such as demosaicing, color correction and gamma correction. These processes can amplify, attenuate or distort the noise in the raw image [42]. Demosaicing is a color interpolation algorithm that is required in some cameras to obtain high-resolution colors from the available color

33

	Noise type	Term	Origin	Manifestation
Illumination dependent	Shot noise	n_{shot}	Physical nature of photons	Additive temporal and spatial variance
	Photo-response non-uniformity (a type of pattern noise)	n_{PRNU}	Imperfections in sensor	Multiplicative spatial variance
Illumination independent	Dark current noise	n_{dark}	Free electrons generated by thermal energy in sensor	Additive temporal and spatial variance
	Readout noise	n_{read}	Noise in circuit that transforms charge from sensor into voltage	Additive temporal and spatial variance
	Fixed-pattern noise (a type of pattern noise)	n_{FP}	Imperfections in sensor	Additive spatial variance
	Quantization noise	n_q	Analog-to-digital conversion of signal	Additive, dependent on image content

Table 2.1: Summary of the noise terms present in raw images captured by CCD or CMOS digital cameras. Information adapted from [43, 44].

channels. It typically has the effect of attenuating spatial and temporal noise variance [43], but introduces spatial correlations [46]. Gamma correction is a nonlinear encoding that is used primarily to reduce the visual artifacts introduced by quantization [47]. The process can be described with a function $g(\cdot)$ that modifies an intensity u according to $g(u) = u^\gamma$. In digital cameras, we often see $0 < \gamma < 1$ [45]; this has the effect of increasing the noise in dark regions and decreasing the noise in bright regions, which we will see later in this chapter.

When the digital image is written to the camera's memory device, it may be compressed, which introduces compression artifacts [42]. While most digital cameras may be configured to output a RAW image file (that is, the image without any post-capture processing effects), this option is often not available for videos.

In the video denoising and image denoising literatures, the noise caused by the acquisition step is commonly assumed to be additive zero-mean white Gaussian noise, even though some noise effects may actually be multiplicative or spatially correlated [43, 48, 49]. Furthermore, all imaging sensors are affected by shot noise, which is

strongly dependent on the amount of incident light. CCD noise models typically assume that the noise is independent between pixels, but is signal dependent in that the noise level at any given pixel depends on the uncorrupted pixel intensity [45]. CMOS image sensors have also been shown to introduce noise with levels that are dependent on the incoming photocurrent [50]. The relationship between the noise level and mean pixel intensity is described by a noise level function (NLF), which is dependent on the natural behavior of photons, the properties of the camera and some recording parameters [34, 51].

2.2.2 SNR estimation

We examine the individual problems of noise and signal estimation. Noise estimation is useful in many computer vision algorithms such as denoising and motion estimation [34]. The general principle of noise estimation is to measure the deviations in intensity over structure-free areas (that is, areas in the image where the signal is known to be homogenous and all variations are caused by noise rather than textures or lighting in the image scene) [52], or over multiple images of the same scene [41]. When one has a multiple images of a static scene containing a wide range of pixel intensities, such as in a video of a stationary color calibration target, the NLF may be estimated by simply measuring the standard deviation over time of the pixels at every mean pixel intensity [41, 51]. Noise estimation from a single image or from videos of non-stationary scenes is more difficult. Noise estimation from a single image tends to work in two ways: by filtering the image to reduce structure and then measuring the noise level from the filtered image [34, 52], or by classifying certain image regions as homogenous and measuring the noise level in those regions [53, 52].

The signal in most traditional image processing literature is defined as the original image before it has been corrupted by noise. The analog to our signal estimation problem in traditional image processing is the problem of image and video denoising, which aims to recover the uncorrupted image from a noisy image. A large class of image denoising techniques rely on the concept of averaging to reduce noise levels. In magnetic resonance imaging, for example, the signal is commonly approximated

using the mean intensity of some user-defined regions of interest [54, 55]. One of the simplest automatic image denoising methods is the Gaussian smoothing approach, in which the the corrupted image is convolved with a Gaussian kernel to reduce the level of Gaussian noise [56]. A more complex local filtering method is anisotropic filtering, which reduces the unwanted blurring effects of the Gaussian kernel by preserving sharp edges [57]. There are many other techniques that use local filtering, frequency domain filtering or non-local filtering [56]. Video denoising techniques use similar techniques to image denoising, but also utilize the temporal dimension to achieve better results. Many video denoising algorithms also apply motion compensation methods in order to preserve temporal coherence [45].

Several key differences make it difficult to directly apply traditional denoising techniques to signal estimation and denoising in video magnification. Firstly, video magnification is affected by ancillary signal and artifactual noise, which may be harder to model than acquisition noise. Many state-of-the art video denoising techniques are capable of removing the types of noise commonly introduced by imaging devices such as additive, multiplicative or structural noise [58, 48], but may not be applicable to more complex noise models. Secondly, video magnification provides the most useful results when applied to videos with low input signal levels. Most existing denoising algorithms focus on applications with relatively high input SNR's, and may not be accurate enough to estimate a low amplitude signal. For example, many video denoising algorithms are evaluated on image sequences with peak signal-to-noise ratio (PSNR) values in the 10-30 dB range, and obtain improvements in PSNR on the order of 0-10 dB [59, 60]. In contrast, in order to visualize blood flow in the human face using Eulerian video magnification, a magnification factor of 100 has been reported [1]; this indicates that the amplitude of the color variation in the input video is low relative to the image intensity, and that the SNR is frequently much lower than 10 dB.

We also consider existing techniques for one-dimensional signals. The problem of speech enhancement bears promising similarities to the problems in video magnification. Speech enhancement aims to produce an uncorrupted speech signal from an

audio segment containing noisy speech. The noise in speech recordings is typically assumed to be additive and signal-independent [61]. Practical speech enhancement algorithms are expected to be effective in real-world environments with low SNR values, e.g., -5dB to 15dB [62]. Speech enhancement differs from most problems in video magnification in that the desired speech signal occupies a wide frequency range (speech signals are often band-limited at $0-4$ kHz [63]). Furthermore, human perception of speech is affected by many spectral features such as the low-pass shape of the speech spectrum, and spectral peaks [62, 64]. As such, most speech enhancement algorithms are concerned with estimating the noise spectrum and recovering the spectrum of the clean speech signal [62]. Speech enhancement algorithms fall into several categories. Spectral subtractive algorithms involve estimating the noise spectrum (typically in the Fourier domain) from a speech-free segment of a recording, and then subtracting the noise spectrum from noisy speech in order to obtain an estimate of the uncorrupted speech signal [62, 65]. Spectral subtractive algorithms are simple to implement. Subspace algorithms use matrix factorization techniques such as singular value decomposition to decompose the speech recording into a subspace that primarily contains the uncorrupted signal and a subspace that primarily contains the noise [62]. Another class of speech enhancement algorithms use statistical models of speech to estimate parameters of the uncorrupted speech signal and noise [66, 62].

In the following sections, we provide definitions of signal and noise suitable for video magnification. We then investigate two algorithms for estimating the signal and noise level of each pixel in a video sequence.

2.3 Signal and noise model

In this section, we define a model to describe the signal and noise in the context of video magnification. This model will allow us to develop algorithms for computing the input SNR.

2.3.1 General case

In Fig. 2-2, we defined the signal of interest to be the variations caused by the "interesting" changes in the video effected by some source of interest. This signal is corrupted by ancillary signal, artifactual noise and acquisition noise. In the Eulerian video magnification work, Wu textitet al. present a model for a video in which the signal is the variations in the video that are caused by rigid motion in the video scene. The rigid motion is effected by sources such as the flow of blood in a human's head, or the opening and closing of a camera shutter [1]. Here, we present a similar model that includes a more general definition of signal, as well the aforementioned noise terms.

Let $I(\boldsymbol{x}, t)$ denote the image intensity at the 2D location $\boldsymbol{x}$ and time t. The average frame of the image sequence is $I_0(\boldsymbol{x})$. We assume that the temporal variations in the video are small relative to the intensity of the mean frame.

We use the term $n_g(\boldsymbol{x}, t)$ to describe acquisition noise, and assume that it may be modeled by additive white Gaussian noise with pixel-dependent variance. We use the term $\phi(\boldsymbol{x}, t, I_0)$ to describe variations caused by the source of interest, which encompasses the signal of interest and the ancillary signal. We describe variations caused by all other sources, or artifactual noise (see Fig. 2-2), using $n_a(\boldsymbol{x}, t, I_0)$. We model the image sequence as follows:

$$I(\boldsymbol{x}, t) = I_0(\boldsymbol{x}) + \phi(\boldsymbol{x}, t, I_0) + n_a(\boldsymbol{x}, t, I_0) + n_g(\boldsymbol{x}, t) \tag{2.1}$$

For now, let us assume that the source of interest only produces one type of change (e.g., a color change with no motion) in the scene. Ideally, video magnification should produce the following:

$$I_{mag}(\boldsymbol{x}, t) = I_0(\boldsymbol{x}) + (1 + \alpha)\phi(\boldsymbol{x}, t, I_0) \tag{2.2}$$

where α is the magnification factor. In practice, some amount of noise is ultimately

also amplified:

$$\hat{I}_{mag}(\boldsymbol{x}, t) = I_0(\boldsymbol{x}) + (1 + \alpha)\big(\phi\left(\boldsymbol{x}, t, I_0\right) + \xi_a n_a\left(\boldsymbol{x}, t, I_0\right) + \xi_g n_g\left(\boldsymbol{x}, t\right)\big)$$

where ξ_a and ξ_g represent some noise reduction factors that may be achieved through a variety of techniques including image stabilization, spatial filtering, and temporal filtering.

2.3.2   Visualizing blood flow

In the specific application of visualizing blood flow, we make additional assumptions that allow us to simplify the model presented in the previous section.

We define the signal to be the intensity variations caused by the color changes in the skin that occur as a result of pulsatile blood flow. In [30], Balakrishnan textitet al. showed that the flow of blood in the head also appears to produce a mechanical effect that causes the head to move, which can be visualized using video magnification. In our work, we define any intensity variations that occur as a result of the mechanical effect of blood flow (e.g. rigid motions of the limb, or pulsating motions visible on the surface of the skin) to be ancillary signal. Equation 2.1 becomes:

$$I(\boldsymbol{x}, t) = I_0(\boldsymbol{x}) + \phi_{color}(\boldsymbol{x}, t, I_0) + \phi_{motion}(\boldsymbol{x}, t, I_0) + \phi_a(\boldsymbol{x}, t, I_0) + n_g(\boldsymbol{x}, t) \qquad (2.3)$$

For now, we assume that the ϕ_{motion} term is negligible after mechanically and/or computationally stabilizing the video subject and image sequence respectively. We will discuss this assumption later. Equation 2.3 simplifies to:

$$I(\boldsymbol{x}, t) = I_0(\boldsymbol{x}) + \phi_{color}(\boldsymbol{x}, t, I_0) + n_g(\boldsymbol{x}, t) \qquad (2.4)$$

We assume that $\phi_{color}(\boldsymbol{x}, t)$ has zero mean. Furthermore, we assume that the amplitude of $\phi_{color}(\boldsymbol{x}, t)$ is small enough that the variance of $n_g(\boldsymbol{x}, t)$ depends only on the

mean image intensity; that is,

$$n(\boldsymbol{x}, t) \sim N(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 & \sigma_n^2(I_0(\boldsymbol{x})) \\ \sigma_n^2(I_0(\boldsymbol{x})) & 0 \end{pmatrix})$$

Henceforth we will use $\phi(\boldsymbol{x}, t)$ to describe our signal of interest, and $n(\boldsymbol{x}, t)$ to describe the acquisition noise.

In the rest of this chapter, we present several approaches for measuring $\sigma_n(I_0(\boldsymbol{x}))$ and the amplitude of $\phi(\boldsymbol{x}, t)$ in order to estimate the input SNR. We discuss the denoising techniques that can be used to reduce $\sigma_n(I_0(\boldsymbol{x}))$ and to obtain an adequate output SNR.

2.4 SNR estimation algorithms

As defined in our model, a video intended for visualizing blood flow consists of the signal of interest and acquisition noise. Based on this model, we present a process for estimating the input SNR. In the following sections, we demonstrate a technique for measuring the acquisition noise levels, and develop two algorithms for estimating the amplitude of the signal of interest.

2.4.1 Noise estimation

As we discussed in section 2.2.1, in an image acquired by a digital camera, the noise level at each pixel depends on the mean pixel intensity. This relationship is described by a noise level function (NLF). The NLF can be estimated from an image as demonstrated by Liu textitet al. [34], or it can be measured from a video of a static scene that contains a range of pixel intensity values. For example, we measured the NLF for a Panasonic Lumix DMC-GF2 using an approach similar to the one described in [51].

We took a video recording of a color calibration target and measured the standard deviation of the green channel intensity over time at each pixel. In this work, we are

only concerned with the green channel, but this algorithm could easily be extended to determine the NLF of other color channels. We generated an NLF (Fig. 2-3) that is relatively flat over the range of possible mean pixel intensities, and resembles one of the NLF's presented in [51]. The shape of an NLF can be affected by post-capture processes in the camera, such as gamma correction. The Lumix G2 supports the sRGB and Adobe RGB color spaces. While the color space used in video recordings is not explicitly mentioned in the camera specifications, both sRGB and Adobe RGB require the application of a gamma correction function that may be approximated as $g(u) = u^{\frac{1}{\gamma}}$ where u is the input image intensity and $\gamma = 2.2$ [67, 68]. This has the effect of increasing the noise level in dark parts of the image and decreasing the noise level in bright parts. Other processes such as tone scale or color correction may have also influenced the shape of the NLF. In our work, we assume that these processes do not introduce any additional noise terms to our image.
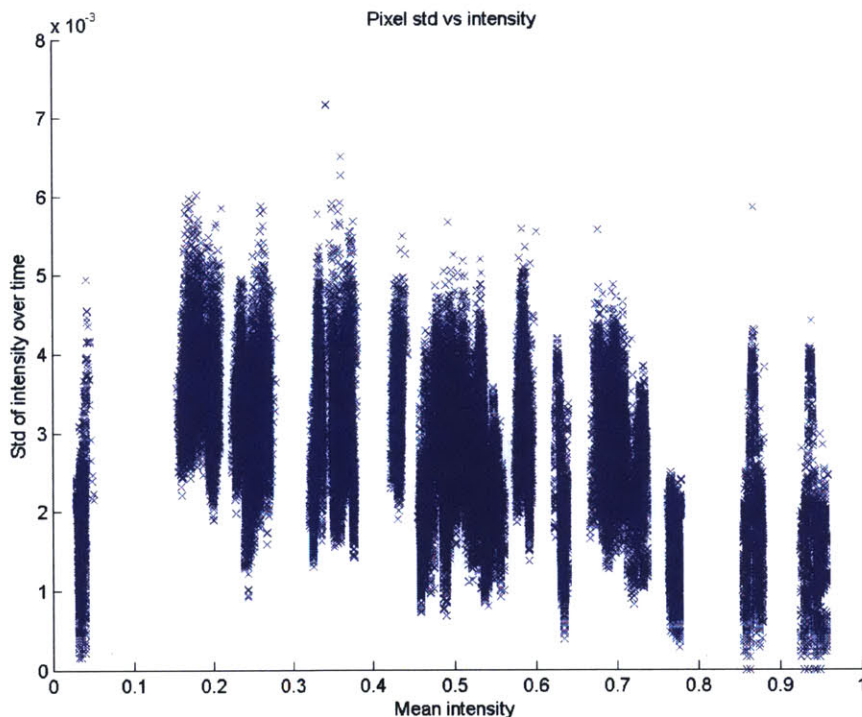
Figure 2-3: The green channel NLF of a Panasonic Lumix DMC-G2, computed by measuring the standard deviation of the green channel intensity of each pixel in a Macbeth Color Checker Card over 600 frames.

2.4.2 Signal amplitude estimation

The amplitude of the signal of interest is more difficult to obtain than acquisition noise levels, particularly when the SNR is low. We discuss two algorithms that use denoising techniques to estimate the amplitude of this signal.

Spectral subtraction

Spectral subtraction is an audio denoising technique that estimates a clean signal from a noise-corrupted acoustical recording by subtracting an estimate of the noise. In this section, we modify the algorithm and use it to denoise a video and produce an estimate of the uncorrupted signal of interest.

The algorithm described in [69] works by estimating the noise magnitude spectrum from a segment of an acoustical recording with no speech activity. This estimated noise magnitude spectrum is then subtracted from the magnitude spectrum of a segment of noisy speech, while maintaining the phase of the noisy speech. The noise is assumed to be additive and signal-independent, but potentially nonstationary (that is, its spectral magnitude may vary slowly over time). Because of the time-varying nature of the noise spectrum, this algorithm is performed over short overlapping segments of the input signal.

We use a modified version of the spectral subtraction algorithm to remove the noise at each pixel. We assume the noise is additive and white, with an amplitude that depends on the mean intensity of the pixel. We assume that the noise levels do not change over time, so we do not segment the signal into short overlapping segments as described in [69]. In the context of video magnification, pixels that are known to contain no signal (e.g., pixels in the background) are analogous to non-speech segments, while pixels within the video subject (e.g., the limb in which we wish to visualize blood flow) are analogous to noisy speech segments. The intensity time series at individual pixels are analogous to the short overlapping input segments used by Boll textitet al. [69]. Our algorithm is as follows:

Spectral subtraction algorithm

1. Identify a region of pixels that contain no signal; we call this the noise region. Ideally, a color calibration target should be placed next to the subject of the video for this purpose. This region should contain as wide a range of mean pixel intensities as possible. We assume that the pixels in this region obey the following:

$$I(\boldsymbol{x}, t) = I_0(\boldsymbol{x}) + n(\boldsymbol{x}, t)$$

2. Identify a region of pixels that contain some noise-corrupted signal; we call this the signal region. We assume that each of the pixels in this region can be represented by the following:

$$I(\boldsymbol{x}, t) = I_0(\boldsymbol{x}) + \phi(\boldsymbol{x}, t) + n(\boldsymbol{x}, t)$$

Our goal is to produce an estimate of the signal $\phi(\boldsymbol{x}, t)$, which we abbreviate as $\hat{s}(t)$.

3. For each pixel in the signal region,

   (a) Compute the mean intensity over time and the Fourier transform of the intensity time series. We use $y(t)$ to denote the intensity time series of this pixel. We use the notation $\mathcal{F}$ to denote the Fourier transform.

   $$y(e^{j\omega}) = \mathcal{F}[y(t)]$$

   (b) Identify N pixels in the noise region that have the closest mean intensity over time. For simplicity, we use the notation $n_i(t)$ where $i = 1, \cdots, P$ to denote the intensity time series of these pixels.

   (c) Compute the mean magnitude spectrum of the N pixels from the noise region. We use the term $\mu(e^{j\omega})$ to describe this value.

   $$\mu(e^{j\omega}) = \frac{1}{P} \sum_{i=1}^{P} |n_i(e^{j\omega})|$$

(d) Subtract the mean magnitude spectrum from the magnitude spectrum of $y(e^{j\omega})$ to estimate the magnitude of $s(e^{j\omega})$.

$$|\hat{s}(e^{j\omega})| = |y(e^{j\omega})| - \mu(e^{j\omega})$$

Any resulting negative magnitude values are set to 0, as described in [69].

(e) Compute the inverse Fourier transform of the frequency spectrum of p_i.

$$\hat{s}(t) = \mathcal{F}^{-1}\left[\left(|y(e^{j\omega})| - \mu(e^{j\omega})\right) e^{j\angle y(e^{j\omega})}\right]$$

(f) Compute the variance of $\hat{s}(t)$ over time.

We omit the residual noise reduction step described in [69], in which the estimated magnitude spectrum after step 3d (above) is adjusted to be more similar to the magnitude spectra of the two segments preceding and following the current segment. This achieves a similar effect to averaging the estimated magnitude spectrum over three segments. In our algorithm, the analogous step would be to adjust the estimated magnitude spectrum at each pixel to be similar to the spectra of the pixels in a neighborhood around the current pixel; this is similar to performing a spatial averaging operation on the estimated magnitude spectrum at each pixel. We found that even using a small averaging neighborhood produced unfavorably smoothed results.

We evaluate the performance of this algorithm on a simulated image sequence in section 2.5.1.

Spatial averaging

Local spatial averaging is used in a large class of image denoising algorithms to reduce noise levels [56]. We present an algorithm that uses spatial averaging in a similar way to estimate the amplitude of the signal of interest in our model. Our algorithm examines the effect of spatial averaging on the noise level in a video, and extrapolates this relationship to estimate the signal power in the presence of no noise.

44

In section 2.3, we defined the signal and noise model:

$$I(\boldsymbol{x}, t) = I_0(\boldsymbol{x}) + \phi(\boldsymbol{x}, t) + n(\boldsymbol{x}, t)$$

Under the assumption that the amplitude of $\phi(\boldsymbol{x}, t)$ is small enough to not affect the value of σ_n, all of the above terms are independent in time and the variance over time of these terms are additive. We use the notation $\text{Var}_t(f(\boldsymbol{x}, t))$ to denote $\text{Var}[f(\boldsymbol{x}, t)|t]$, or the variance of the function $f(\boldsymbol{x}, t)$ taken over the time dimension.

$$\text{Var}_t[I(\boldsymbol{x}, t)] = \text{Var}_t[I_0(\boldsymbol{x})] + \text{Var}_t[\phi(\boldsymbol{x}, t)] + \text{Var}_t[n(\boldsymbol{x}, t)]$$

Or, simply:

$$\sigma_I^2(\boldsymbol{x}) = \sigma_s^2(\boldsymbol{x}) + \sigma_n^2(\boldsymbol{x}) \tag{2.5}$$

Throughout this work, we will refer to the variance of image intensity over time, $\sigma_I^2(\boldsymbol{x})$, as the **total variance** since it represents the sum of the signal variance and noise variance. When we work with real videos later in this chapter, we will refer to this value as the **total sample variance**.

In general, applying a spatial averaging filter to an image with additive, signal-independent zero-mean Gaussian noise will attenuate the noise variance by a factor ξ that is inversely proportional to the size of the filter. For instance, an averaging filter (or, box filter) of size $m \times m$ attenuates the noise by a factor of $\xi = \frac{1}{m^2}$ [70], while a Gaussian filter with standard deviation σ_f produces a factor of $\xi \propto \dfrac{1}{8\pi\sigma_f^2}$ [45]. In the context of video magnification, if we convolve each frame with a box filter of size $m \times m$, the resulting noise variance $\sigma_{n,m}^2$ is described by:

$$\sigma_{n,m}^2 = \frac{\sigma_n^2(\boldsymbol{x})}{m^2}$$

For now, we assume that the signal $\phi(\boldsymbol{x}, t)$ does not vary over space, that is, $\phi(\boldsymbol{x}, t) = \phi(t)$ for all $\boldsymbol{x}$. We will relax this constraint later. The resulting variance of

the image after convolution with a box filter, $\sigma_{I,m}^2(\boldsymbol{x})$, is described by:

$$\sigma_{I,m}^2(\boldsymbol{x}) = \sigma_s^2 + \frac{\sigma_n^2(\boldsymbol{x})}{m^2} \tag{2.6}$$

The proof of Eq. 2.6 is fairly simple. Let $I_m(\boldsymbol{x}, t)$ represent the intensity time series of the pixel at location $\boldsymbol{x}$ in an image that has been convolved with an $m \times m$ filter.

$$I_m(\boldsymbol{x}, t) = \frac{1}{m^2} \sum_{i=1}^{m^2} I(\boldsymbol{x_i}, t)$$

where each $\boldsymbol{x}_i$ is a location within the filter convolution neighborhood surrounding the pixel.

$$I_m(\boldsymbol{x}, t) = \frac{1}{m^2} \sum_{i=1}^{m^2} \big(I_0(\boldsymbol{x_i}) + \phi(\boldsymbol{x_i}, t) + n(\boldsymbol{x_i}, t) \big)$$

$$\mathrm{Var}_t[I_m(\boldsymbol{x}, t)] = \mathrm{Var}_t[\frac{1}{m^2} \sum_{i=1}^{m^2} [I_0(\boldsymbol{x_i})] + \mathrm{Var}_t[\frac{1}{m^2} \sum_{i=1}^{m^2} [\phi(t)]_t + \mathrm{Var}_t[\frac{1}{m^2} \sum_{i=1}^{m^2} [n(\boldsymbol{x_i}, t)]$$

$$= \mathrm{Var}_t[\phi(t)] + \frac{1}{m^4} \sum_{i=1}^{m^2} \mathrm{Var}[n(\boldsymbol{x_i}, t)]_t$$

$$\sigma_{I,m}^2(\boldsymbol{x}) = \sigma_s^2 + \frac{\sigma_n^2(\boldsymbol{x})}{m^2}$$

This proof still holds if we assume, rather than a spatially invariant signal $\phi(t)$, that there is some average signal at location $\boldsymbol{x}$ after the convolution of the image with a filter of size $m \times m$. We write this average signal term as:

$$\phi_m(\boldsymbol{x}, t) = \frac{1}{m^2} \sum_{i=1}^{m^2} \phi(\boldsymbol{x_i}, t)$$

$$\mathrm{Var}_t[\phi_m(\boldsymbol{x}, t)] = \sigma_{s,m}^2(\boldsymbol{x})$$

Equation 2.6 becomes:

$$\sigma_{I,m}^2(\boldsymbol{x}) = \sigma_{s,m}^2(\boldsymbol{x}) + \frac{\sigma_n^2(\boldsymbol{x})}{m^2} \tag{2.7}$$

46

Later in this section, we will describe an algorithm that estimates the value of $\sigma_{s,m}^2(\boldsymbol{x})$. It is important to note that $\sigma_{s,m}^2(\boldsymbol{x})$ is a good approximation of the variance of the true signal over time $\sigma_s^2(\boldsymbol{x})$ only for sufficiently small m; if m is too large, the blurring effect distorts the estimated signal. The appropriate value of m depends on several factors including the spatial nature of the signal (which depends on the phenomenon of interest) and video resolution. In the context of visualizing blood flow, we typically assume that the signal varies gradually enough in space that a fairly large m can be used; we discuss this in the next chapter.

What we have presented thus far is analogous to the use of spatial averaging in image denoising. Several image denoising algorithms, such as Gaussian smoothing and the anisotropic filtering, use a Gaussian filter to reduce the variance of additive white Gaussian noise in an image to produce an estimate of uncorrupted image [57]. Say an image $I(\boldsymbol{x})$ is corrupted by additive white Gaussian noise as follows:

$$I(\boldsymbol{x}) = U(\boldsymbol{x}) + n(\boldsymbol{x})$$

where $U(\boldsymbol{x})$ is the uncorrupted image (which is the signal of interest), and $n(\boldsymbol{x}) \sim N(0, \sigma_n^2(U(\boldsymbol{x})))$. Applying a Gaussian filter of variance σ_f^2 produces the smoothed image:

$$I_m(\boldsymbol{x}) = U_m(\boldsymbol{x}) + n_m(\boldsymbol{x})$$

where $\mathrm{Var}[n_m(\boldsymbol{x})] = \xi \sigma_n^2$ and $\xi \propto \dfrac{1}{8\pi\sigma_f^2}$ [45].

The smoothed image $I_m(\boldsymbol{x})$ is a reasonable estimate of $U_m(\boldsymbol{x})$, when the smoothed noise level is low compared to the amplitude of the uncorrupted image; that is, $I_m(\boldsymbol{x}) \simeq U_m(\boldsymbol{x})$ if $\xi \sigma_n^2 \ll |U(\boldsymbol{x})|^2$. In parts of the image where the signal $U(\boldsymbol{x})$ is spatially invariant, $U_m(\boldsymbol{x}) \simeq U(\boldsymbol{x})$. So we see that for $\xi \sigma_n^2 \ll |U(\boldsymbol{x})|^2$, and sufficiently small m (where the meaning of "sufficient" depends on the signal content), $I_m(\boldsymbol{x}) \simeq U(\boldsymbol{x})$. As we discussed in section 2.2.2, the noise levels in images are typically small relative to the signal amplitude, so the assumption of $\xi \sigma_n^2 \ll |U(\boldsymbol{x})|^2$ is

valid.

In the context of video magnification, however, we hypothesize that our signal amplitude is similar to the noise level even after the noise has been attenuated using a spatial averaging filter. That is, $\sigma_s^2(\boldsymbol{x}) \ll \sigma_n^2(\boldsymbol{x})$ and $\sigma_{s,m}^2(\boldsymbol{x}) \leq \frac{\sigma_n^2(\boldsymbol{x})}{m^2}$ for most values of $\boldsymbol{x}$. To estimate a signal amplitude that is small relative to the noise amplitude, we propose to estimate $\sigma_m^2(\boldsymbol{x})$ for $\frac{\sigma_n^2(\boldsymbol{x})}{m^2} = 0$, that is, for $m = \infty$. Our algorithm is as follows:

Spatial averaging algorithm

1. Select a range of $m_i \times m_i$ box filter sizes, for instance, $m_1 = 3, m_2 = 5, \cdots, m_F = 9$. m_F should be chosen such that $\sigma_{s,m_F}^2(\boldsymbol{x}) \simeq \sigma_s^2(\boldsymbol{x})$.

2. For each filter size m_i, where $i = 1, 2, \cdots, F$:

   (a) Apply the filter using a 2D convolution to each frame of the image sequence.

   (b) Measure $\text{Var}_t[I_{m_i}(\boldsymbol{x}, t)]$, that is, the sample variance of the intensity over time at pixel location $\boldsymbol{x}$.

3. Define $u_i = \frac{1}{m_i^2}$. For each pixel, compute a least squares regression line $\text{Var}_t[I_{m_i}(\boldsymbol{x}, t)] = \hat{\beta}_0(\boldsymbol{x}) + \hat{\beta}_1(\boldsymbol{x}) \cdot u$, where $\beta_0(\boldsymbol{x}) = \sigma_s^2(\boldsymbol{x})$ and $\beta_1(\boldsymbol{x}) = \sigma_n^2(\boldsymbol{x})$.

This algorithm is summarized in Fig. 2-4. The algorithm gives an estimate of the variance of the signal over time, $\sigma_s^2(\boldsymbol{x})$, which, under the assumption that $\phi(\boldsymbol{x}, t)$ has zero mean, is equal to the signal power P_ϕ. The algorithm also produces an estimate of noise variance $\sigma_n^2(\boldsymbol{x})$.

In the following section, we evaluate the accuracy of these estimates using a simulated image sequence. Later in this chapter, we evaluate this algorithm on a real video of a hand.

2.5 SNR estimation in a simulated video

To test our signal estimation algorithms, we simulated a periodic color change that is corrupted by acquisition noise. The image sequences were created and then analyzed
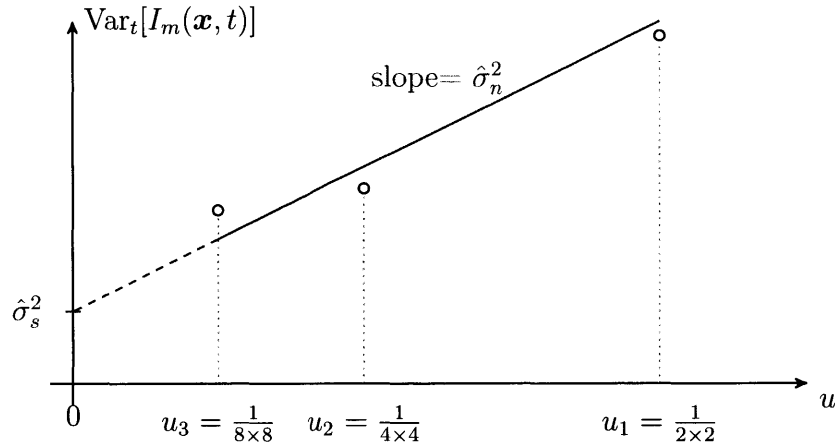
Figure 2-4: Summary of the spatial averaging algorithm. We first estimate the fit of $\text{Var}_t[I_m(\boldsymbol{x}, t)]$ (the total variance) against u (the inverse filter area, not shown to scale in the figure). The slope of the fit estimates the noise variance (the component of the total variance that is affected by spatial averaging). The intercept estimates the signal variance (the component of the total variance that is unaffected by spatial averaging).

within the same MATLAB script to avoid the quantization errors that would result from generating and reading a video file. We used a 64-bit representation for the pixel intensity values such that each pixel takes on a value within the range $[0, 1]$. The general format of each image sequence is described below.

Each video frame is split into two square regions. Region A contains only noise, while region B contains a signal as well as noise, that is:

$$I_A(\boldsymbol{x}, t) = I_0(\boldsymbol{x}) + n(\boldsymbol{x}, t)$$

$$I_B(\boldsymbol{x}, t) = I_0(\boldsymbol{x}) + \phi(\boldsymbol{x}, t) + n(\boldsymbol{x}, t)$$

We defined our mean image to be spatially invariant, our signal to be a spatially invariant sinusoid, and our noise to be additive zero-mean Gaussian noise, as follows:

$$I_0(\boldsymbol{x}) = I_0$$

$$\phi(\boldsymbol{x}, t) = A_s \sin(2\pi t f / F_s)$$

49

$$n(\boldsymbol{x}, t) \sim N(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 & \sigma_n^2 \\ \sigma_n^2 & 0 \end{pmatrix})$$

The choice of $\phi(\boldsymbol{x}, t) = A_s \sin(2\pi t f_{HR}/F_s)$ calls for some explanation. The color change in the skin that is caused by blood flow is not likely to be purely sinusoidal, given that blood volume waveforms obtained via photoplethysmography are not sinusoidal [28]. However, the color change is likely to have some frequency component at the heart rate f. We selected a sine wave to be a simplistic representation of a portion of the actual signal.

To test the performance of our signal estimation algorithms on different signal amplitudes, we created two image sequences using $A_s = 0.1$ and $A_s = 0.01$. All other parameter values were held constant: $\sigma_n = 0.1$, $I_0 = 0.4$, $f = 1.5$, $F_s = 30$. Note that for our zero-mean sinusoidal signal, the power is equal to the variance over time. The variance over time and amplitude are related as follows:

$$\sigma_s^2 = \frac{A_s^2}{2}$$

We investigate the performance of the spectral subtraction algorithm and the spatial averaging algorithm in estimating the signal amplitude in each simulated image sequence. First, we evaluate each algorithm on the simulated image sequence with $A_s = 0.1$.

2.5.1 Spectral subtraction results

We define region A to be the noise region and region B to be the signal region. Using the algorithm defined in section 2.4.2, we produced an estimate of the signal in every pixel within the signal region. We computed the sample variance of the estimated signal at each pixel, $s_s^2(\boldsymbol{x})$, and the mean sample variance over all of the pixels, $\bar{s}_s^2 = \frac{1}{N^2} \sum_{i=1}^{N^2} s_s^2(\boldsymbol{x}_i)$. The mean sample variance is the estimate of the power of the signal of interest. We did this for several values of N, which represents the number of noisy samples from which we derive the mean noise magnitude spectrum. We expected that increasing N would cause the sample mean noise magnitude spectrum

to approach the true mean noise spectrum, which would increase the accuracy of the signal amplitude estimate.
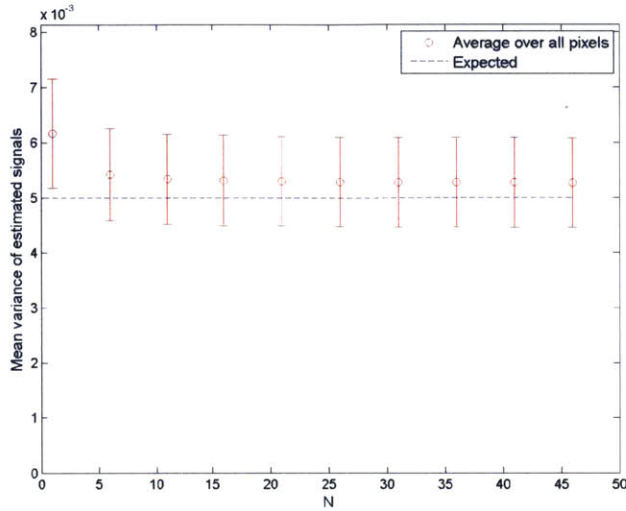
Figure 2-5: The estimated signal power, $\bar{s}_s^2$, produced by applying our spectral subtraction algorithm to a video with $\sigma_n = 0.1$, $A_s = 0.1$. The error bars represent the standard deviation in the values of s_s^2 over all of the pixels in the signal region. $\bar{s}_s^2$ appears to become more accurate as N increases; howver, it appears to converges upon 5.4×10^{-3} rather than the true signal power $\sigma_s^2 = 5.0 \times 10^{-3}$.

In Fig. 2-5 we see that $|\bar{s}_s^2 - \sigma_s^2|$ decreases as N increases, as we expected. However, $\bar{s}_s^2$ converges upon 5.4×10^{-3} rather than $\sigma_s^2 = 5.0 \times 10^{-3}$. This discrepancy is not very large relative to the true signal power, but we will see later in this chapter that this error greatly impacts our estimate of the signal amplitude when the true signal amplitude is very low compared to the noise level. In Fig. 2-6, we see that the estimated signal for $P = 50$ (Fig. 2-6b) is much closer to a sinusoid than the original input (Fig. 2-6a), but still appears to contain some noise.

The spectral subtraction algorithm described by Boll textitet al. is known to produce audible artifacts that are frequently referred to as musical noise [71, 61]. Several authors have attempted to reduce the residual noise through techniques such as over-subtracting the noise spectrum [72], performing non-linear spectral subtraction [73], or optimizing the parameters used in spectral subtraction [71], but we do not investigate these algorithms in this work. In [71], Sim textitet al. describe these artifacts
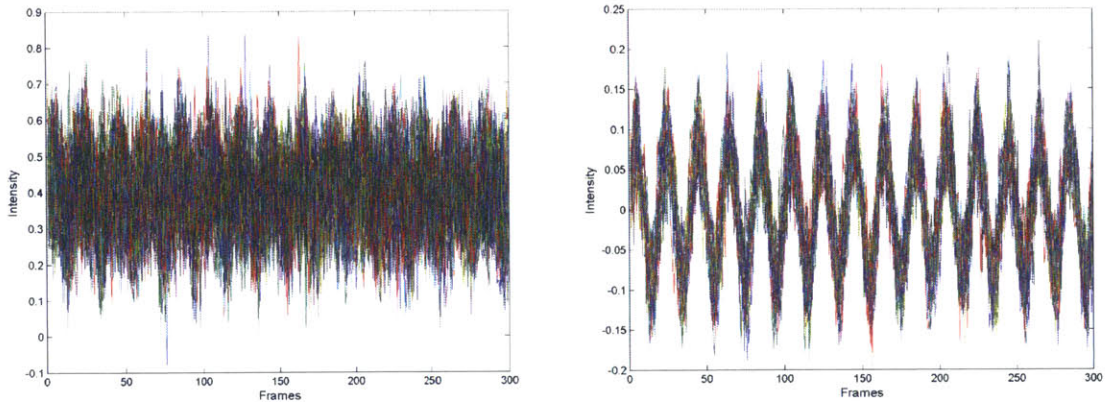
(a) The intensity of the input video $I(\boldsymbol{x}, t)$, plotted for several arbitrary values of $\boldsymbol{x}$.

(b) The estimated signal $\hat{\phi}(\boldsymbol{x}, t)$ produced using $P = 50$, plotted for several arbitrary values of $\boldsymbol{x}$.

Figure 2-6: Input pixel intensities (left) and estimated signal (right) for the simulated video with $A_s = 0.1$, $\sigma_n = 0.1$.

as consisting of two components: (1) remaining noise that has the same perceptual characteristics as the original noise, and (2) additional artifacts that are caused by the presence of random spectral peaks in the estimated spectrum. Fig. 2-6b shows that our estimated signal contains some noise that visually resembles additive white Gaussian noise; however, investigating the nature of the remaining noise is beyond the scope of this work.

2.5.2   Spatial averaging results

We used the spatial averaging algorithm to estimate the signal power in both regions of the video. We applied several $m \times m$ filters, where:

$$m = 3, 5, 7, 9$$

$$u = \frac{1}{m^2} = \frac{1}{3^2}, \frac{1}{5^2}, \frac{1}{7^2}, \frac{1}{9^2}$$

Using the algorithm described in section 2.4.2, we measured the sample variance over time at each pixel, $\mathrm{Var}_t[I_m(\boldsymbol{x}, t)]$, for each value of u. To obtain an accurate mea-
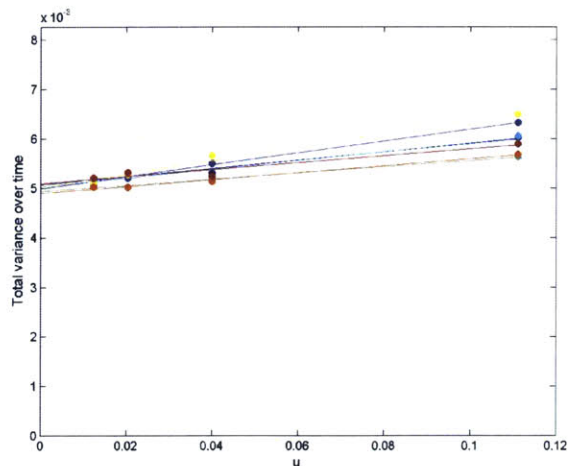
(a) Estimated fit of $\text{Var}_t[I_m(\boldsymbol{x}, t)]$ against u for several of the uncorrelated pixels; each color represents the fit computed from a single pixel. The standard deviation of the slope and intercept values over all the uncorrelated pixels are 3.06×10^{-3} and 1.10×10^{-4} respectively.
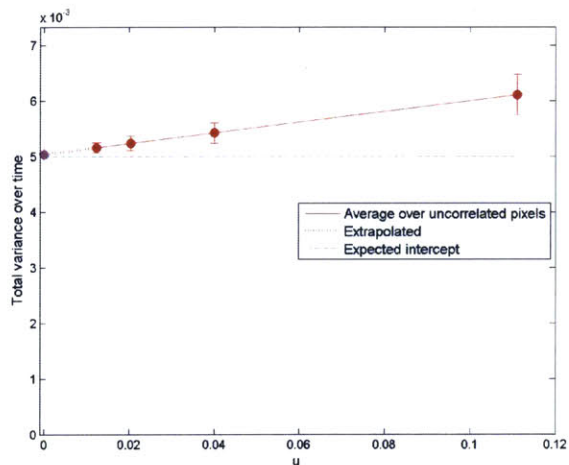
(b) The mean fit curve of $\text{Var}_t[I_{m_i}(\boldsymbol{x}, t)]$ against u. This curve is computed using the mean of the estimated linear regression coefficients produced at each of the uncorrelated pixels. The red error bars represent the standard deviation of $\text{Var}_t[I_{m_i}(\boldsymbol{x}, t)]$ computed over the uncorrelated pixels.

Figure 2-7: Estimated fits of $\text{Var}_t[I_m(\boldsymbol{x}, t)]$ against u for the simulated video in which $A_s = 0.1, \sigma_n = 0.1$. The slope of the fit is an estimate of σ_n^2, and the intercept is an estimate of σ_s^2.

sure of the standard deviation of our results, we selected only spatially uncorrelated pixels; that is, pixels that are greater than 9 px apart. We also excluded pixels that were within 9 px of the image borders to avoid the edge effects of the convolution process. We computed a least squares regression line of $\text{Var}_t[I_m(\boldsymbol{x}, t)]$ against u for

each uncorrelated x. The slope coefficient is an estimate of $\hat{\sigma}_n^2(x)$ and the intercept is an estimate of $\hat{\sigma}_s^2(x)$.

Fig. 2-7 shows that the spatial averaging algorithm produced an accurate estimate of the true signal power from the simulated video in which $A_s = 0.1, \sigma_n = 0.1$. The algorithm produces an estimate of the signal power and noise variance at each pixel, as shown in Fig. 2-7a. There is some variation between the estimated slope and intercept produced at each pixel; however, the standard deviation of these estimates is relatively low. The relatively small standard deviation in this estimate over all pixels suggests that it is even possible to obtain a fairly accurate estimate from a single pixel. To investigate this, we created an estimated signal power map, shown in Fig. 2-8. Due to the simulated white Gaussian noise in the image sequence, these sample standard deviations vary slightly between image sequences created with the same parameter values. Future work should derive the expected values of these deviations.
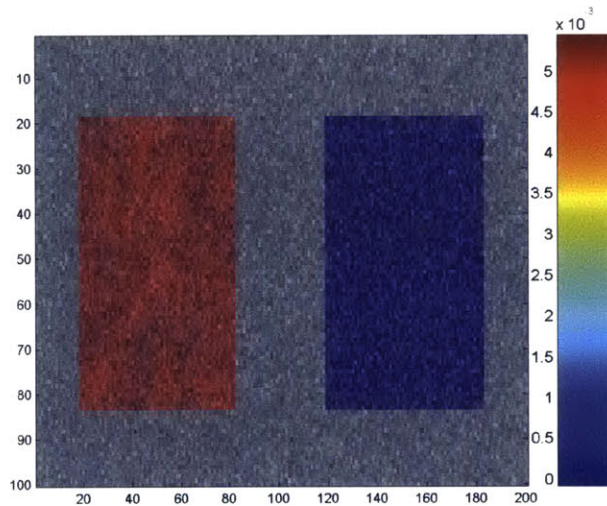
Figure 2-8: Estimated signal power at each pixel in the signal region (left) and noise region (right), for the simulated video in which $\sigma_n = 0.1, A_s = 0.1$. The average estimated signal power in the signal region is 5.02×10^{-3}; the average estimated signal power in the noise region is 3.62×10^{-7}, which is small relative to the estimate from the signal region and can be considered to be close to the true value of 0. The covariance in the estimates of spatially adjacent pixels was measured to be negligible.

Fig. 2-8 shows that the spatial averaging algorithm can be used to produce an

accurate map of the estimated signal power in a video. This visualization is useful for analyzing real videos, which we will see in section 2.6.

2.5.3 Comparison of signal estimation algorithms

In this section, we compare the mean estimated signal power computed by each algorithm. Each algorithm produces the estimate from 289 spatially uncorrelated pixels.

$A_s = 0.1, \sigma_s^2 = 5.00 \times 10^{-3}$	$\hat{\sigma}_s^2$	Standard deviation of $\hat{\sigma}_s^2$
Spectral subtraction algorithm	5.49×10^{-3}	7.91×10^{-4}
Spatial averaging algorithm	4.99×10^{-3}	1.22×10^{-4}

Table 2.2: Estimates of the signal power computed by the spectral subtraction algorithm and spatial averaging algorithm for the simulated video in which $\sigma_n = 0.1, A_s = 0.1$.

$A_s = 0.01, \sigma_s^2 = 5.00 \times 10^{-5}$	$\hat{\sigma}_s^2$	Standard deviation of $\hat{\sigma}_s^2$
Spectral subtraction algorithm	1.29×10^{-3}	2.75×10^{-4}
Spatial averaging algorithm	5.41×10^{-5}	2.61×10^{-5}

Table 2.3: Estimates of the signal power computed by the spectral subtraction algorithm and spatial averaging algorithm for the simulated video in which $\sigma_n = 0.1, A_s = 0.01$. The spectral subtraction algorithm is not able to estimate the true signal power from individual pixels.

The two algorithms produced comparable results for $\sigma_n = 0.1, A_s = 0.1$ (Table 2.2). The lower standard deviation achieved by the spatial averaging algorithm indicates that it can make more accurate estimates from single pixels. The spectral subtraction algorithm fails to produce an accurate estimate of the signal power for $\sigma_n = 0.1, A_s = 0.01$ (Table 2.3). These results indicate that the spatial averaging algorithm is better than the spectral subtracting algorithm at estimating the signal amplitude in videos with low SNRs.

2.6   Applications to visualizing blood flow

In this section, we apply our model and signal estimation algorithm to a real video. We demonstrate the process for estimating the input SNR in a video of a healthy human hand, and evaluate our results. We briefly discuss the effects of the denoising techniques used in Eulerian linear video magnification. We estimate the output SNR that can be achieved by applying these denoising techniques to our video.

2.6.1   Method

We used a Panasonic Lumix DMC-GF2 to take a video of a healthy female subject's hand and wrist. The subject was seated comfortably next to a table that was at approximately the height of her chest. Small '+' marks were drawn on the palm to serve as markers that can be used to estimate the motion of the hand.
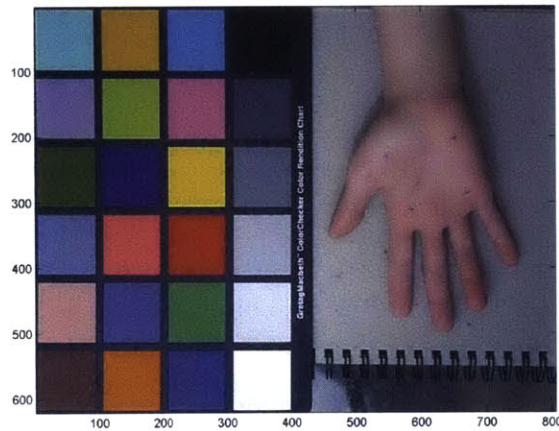
Figure 2-9: A healthy female's hand next to a color calibration target.

We processed the video in MATLAB using a 64-bit representation for each pixel, such that each pixel takes on a value within the range $[0, 1]$. We manually selected a signal region that contained only the palm of the hand, to reduce any artifactual noise that might be caused by the fingers twitching. We selected a noise region that contained the color calibration target. A simple edge detection algorithm was used to exclude any sharp gradients in the hand as well as the borders of the color calibration

56

target. We also applied a simple erosion algorithm to exclude the pixels near the edges of each region that would be affected by the spatial filter.

We applied our spatial averaging algorithm over 450 frames of the video, using $m = 5, 7, 9, 11, 13, 15$, to produce estimates of the signal amplitude and noise variance at each pixel within the selected regions. The values of m were experimentally determined; future work should examine the effect of the choice of m values on the accuracy of the estimates produced by the spatial averaging algorithm.

2.6.2 Noise level estimation

As we demonstrated in section 2.4.1, the noise level in a video can simply be measured from a stationary color calibration target. In this section, we evaluate the noise levels estimated by the spatial averaging algorithm, and compare them to the true noise levels measured from the color calibration target.

Using the spatial averaging algorithm, we estimated the average noise standard deviation in the color calibration target to be 6.10×10^{-3}. The mean standard deviation measured from the color calibration target is 2.63×10^{-3}. Fig. 2-10a shows that for every mean intensity, the noise levels predicted by the spatial averaging algorithm are usually higher than those measured from the video, and had a wider spread. The difference between the mean estimated and mean measured noise levels (Fig. 2-10b) appears to be consistent across every intensity bin. This indicates that there is some systematic error in the spatial averaging algorithm. Future work should investigate the nature of this error.

The noise levels in the color calibration target exhibit a grid-like spatial pattern (Fig. 2-11). The spatial averaging algorithm produces the same pattern as the measured noise levels, although the noise level at each pixel is uniformly higher. This pattern indicates the presence of structured noise, which is common in real videos [48]. Structured noise may be introduced by processes such as demosaicing [46], or H.264 compression [74]. Since structured noise can be difficult to model, future work should endeavor to reduce this type of noise by using a digital camera does not require demosaicing, and is capable of recording raw, uncompressed video files.
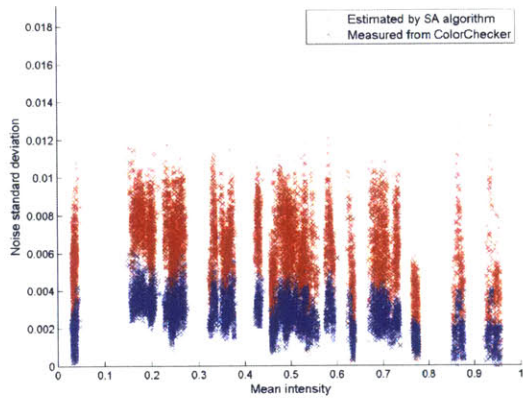
(a) Green channel NLF.
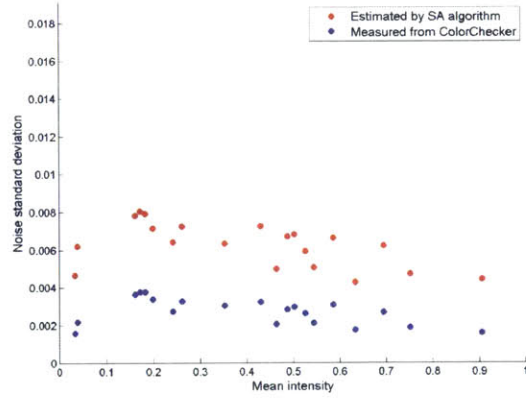
(b) Mean green channel NLF.

Figure 2-10: Using the color calibration target, we measured the noise standard deviation in the green channel. We also used our spatial averaging algorithm to estimate the noise standard deviation in the green channel. The average NLF (right) was computed by binning the noise levels into 20 bins based on mean intensity and then taking the mean noise level within each bin.
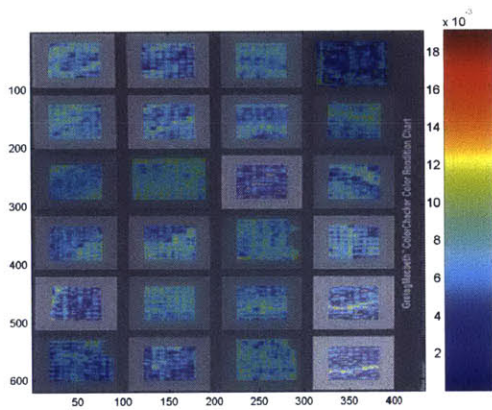
(a) Estimated noise standard deviation.
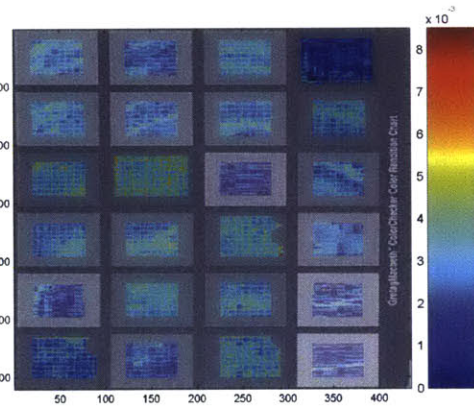
(b) Measured noise standard deviation.

Figure 2-11: Estimated noise standard deviation at each pixel. The estimated noise standard deviation from the spatial averaging algorithm (left) fall into the same grid pattern as the measured noise standard deviation (right).

The average estimated noise standard deviation in the hand is 2.04×10^{-2}; this is substantially higher than the estimated noise standard deviation in the color calibration target. The markers that were drawn on the palm exhibit particularly high noise levels (Fig. 2-12). The high noise levels in the hand are likely to be caused by

small motions. As shown in Fig. 2-4, in the spatial averaging algorithm, we estimate the relationship between the total variance (taken over time) at each pixel and the inverse area of the box filters that are applied to the video. The slope describes the component of the total variance that is affected by the spatial averaging filter. In the algorithm, this component is assumed to consist of just the acquisition noise. However, in a real video, the slope may be artificially inflated if the pixel experiences additional variations that are attenuated by the spatial averaging, such as motion. Small motions can significantly increase the total variance of pixels in areas with high intensity gradients (e.g., near the markers, or in regions with visible texture such as the palm). However, the application of a spatial averaging filter smoothes these intensity gradients and reduces the contribution of motion to the total variance. This is likely to cause the slope of the fit of total variance against inverse filter area to be higher near the markers and in the textured parts of the hand. We discuss the effect of motion further in the next section, where we estimate the signal amplitude in the hand.
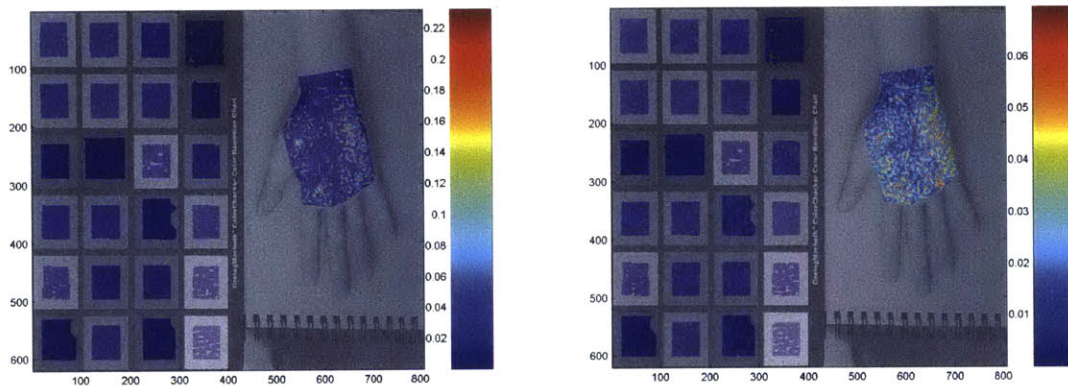
(a) Estimated noise standard deviation.

(b) Estimated noise standard deviation (top 1% of values removed to improve viewing scale).

Figure 2-12: The estimated noise standard deviation is higher in the signal region than in the noise region. Pixels with negative estimated noise variance[1] were excluded.

---

[1]For some pixels, the spatial averaging algorithm produced a negative estimate of the noise variance and/or a negative estimate of the signal power. This is expected to occur with some low probability. The noise in the video inevitably introduces noise to the measured total variance at

2.6.3 Signal amplitude estimation

We used our spatial averaging algorithm to estimate the variance (over time) of the signal in each pixel. We assume that the signal has zero mean, so we refer to this estimated value as the estimated signal power. We refer to the square root of the estimated value as the estimated signal amplitude.
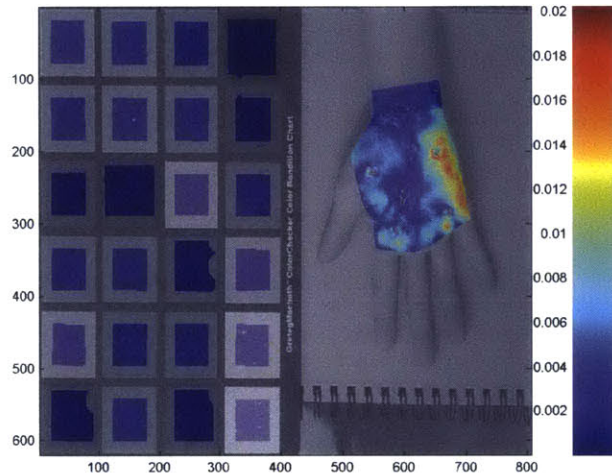
Figure 2-13: The estimated signal amplitude in the pixels of the hand are generally larger than those in the color calibration target.

As expected, the estimated signal amplitudes are low in the color calibration target (with a mean of 7.81×10^{-4}) and higher in the hand (with a mean of 5.32×10^{-3}). The regions with high signal amplitudes appear to agree with the magnified videos we obtain by applying Eulerian linear video magnification to videos of healthy hands (which we discuss in section 3.3). It is unclear if the high signal amplitudes on the right side of the signal region truly indicate strong tissue perfusion in that area. Future work should investigate whether the estimated signal amplitudes reflect the expected blood flow patterns in different scenarios. For instance, if both arteries are occluded with a blood pressure cuff, the estimated signal amplitudes should be low in all parts of the hand.

Fig. 2-14 shows the distribution of the estimated signal amplitudes in the hand.

each pixel, which produces noisy estimates of the noise variance and signal power.
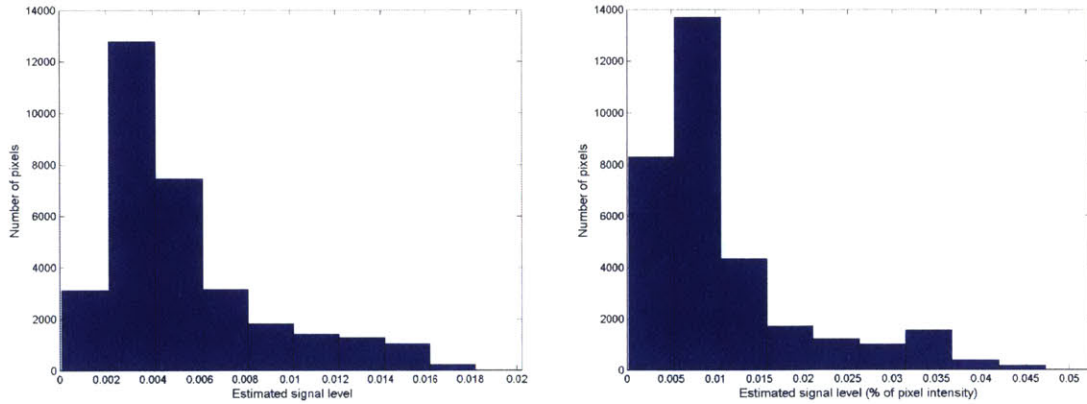
(a) Histogram of signal amplitudes estimated at each pixel in the hand.

(b) Histogram of signal amplitudes estimated at each pixel in the hand, normalized by the mean intensity of the pixel.

Figure 2-14: Distribution of estimated signal amplitudes in the pixels of the hand.

The mean signal amplitude is estimated to be 5.32×10^{-3}, and the mean normalized signal amplitude is estimated to be 1.1% of the corresponding pixel intensity. In section 2.4.1, we noted that the videos produced by our camera were likely to be gamma encoded with an approximate factor of $\frac{1}{2.2}$. After applying the inverse factor to decode our pixel intensity values, we obtain mean estimated signal amplitude of 0.0093% of the irradiance (in the green channel) at each pixel.
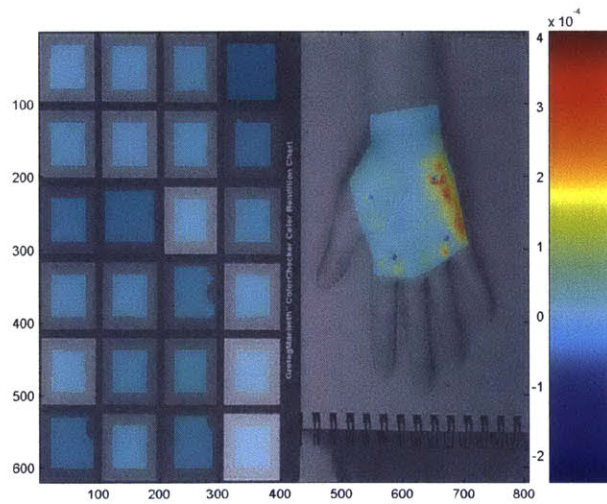
Figure 2-15: Estimated signal power.

61

In Fig. 2-15, we see that pixels near the markers on the hand produced negative estimates of signal power. This was likely caused by the presence of motion. The spatial averaging algorithm produces an estimate of the signal power by first plotting the total variance of a pixel against the inverse area of the spatial averaging filters that are applied. The estimated fit is extrapolated to estimate the total variance when the filter has infinite area (which should reduce the contribution of the noise variance to 0). As we discussed earlier, the presence of motion can increase the slope of this fit, which consequently decreases the extrapolated intercept. Figs. 2-12 and 2-15 suggest that the presence of motion negatively impacts the accuracy of our signal and noise estimates, so it is important for future work to attempt to quantify and remove its effects.

2.6.4   Input SNR

We used the signal power and noise variance estimated in the previous sections to compute an estimate of the input SNR at each pixel in the hand, according to the equation:

$$SNR(\boldsymbol{x}) = 10 \log_{10} \left(\frac{\hat{\sigma}_s^2(\boldsymbol{x})}{\hat{\sigma}_n^2(\boldsymbol{x})} \right)$$

In section 2.6.2, we described several conflicting estimates of the noise levels in the video. The average estimated noise standard deviation computed using the pixels in the hand is 2.04×10^{-2}, which is much higher the mean computed using the color calibration target (6.10×10^{-3}). Both values are higher than the average measured noise standard deviation in the color calibration target (2.63×10^{-3}). The differences between the estimated values in the hand and the color calibration target were likely caused by small motions in the hand. As we discussed earlier, the presence of motion might cause the spatial averaging algorithm to overestimate the noise variance and underestimate the signal power for some pixels. The difference between the estimated and measured noise levels in the color calibration target was likely caused by some systematic error in the spatial averaging algorithm. To compute the SNR at each

62

pixel in the hand, we chose to use the estimated noise variance produced by the spatial averaging algorithm at each pixel. We aimed to produce the most conservative estimate of the input SNR by choosing to use the option with the highest average noise variance.
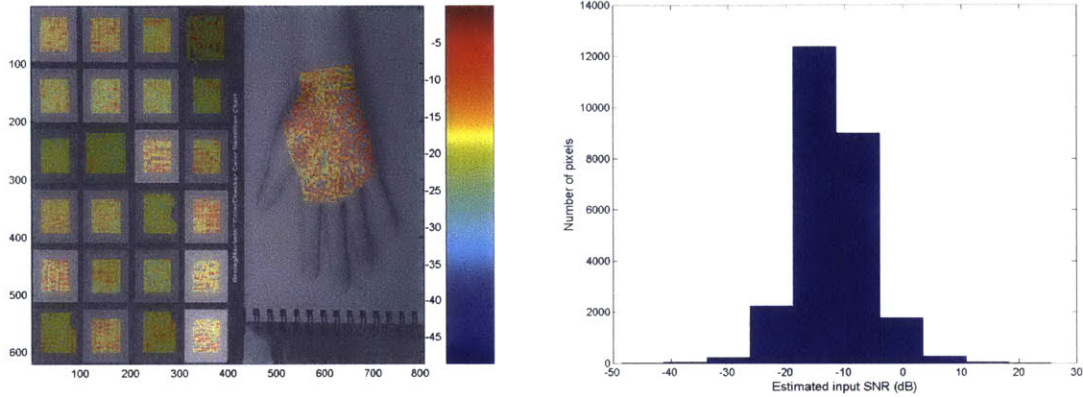
(a) Estimated input SNR (dB) at each pixel. We excluded pixels that had negative estimates of signal power or noise variance, and pixels that had very high or very low estimated SNRs. These excluded pixels appear to be gray in the image.

(b) Histogram of estimated input SNR (dB) from the pixels in the hand.

Figure 2-16: Estimated SNR (dB) in the input video, produced by applying the spatial averaging algorithm to the video.

The mean input SNR in the hand was estimated to be -12.1 dB. This estimate is consistent with our hypothesis that the SNR in the input video is lower than what is typically seen in the image and video processing literature. We also hypothesized that the signal power is much smaller than the noise power. On average, the signal power was estimated to be 11.1 times smaller than the noise power. The SNR map in Fig. 2-16a shows that there is a significant amount of noise in our estimated SNR values. We excluded pixels that produced negative estimates of noise variance or signal power from the SNR map; these pixels are gray in color. The many gray areas in the hand indicate that many pixels produced unreliable estimates. Furthermore, some pixels in the color calibration target appear to have a similar SNR to many parts of the hand.

It is likely that the motion of the hand affected our estimates of the input SNR. While we hypothesized that motion causes the pixels in textured regions to underestimate the SNR, it is unclear how motion affects the pixels in texture-less regions. The reliability of our SNR estimates can be improved by quantifying the effect of motion, as well as other issues mentioned in the previous sections such the structured noise introduced by the camera.

2.6.5 Denoising in Eulerian linear video magnification

In the previous section, we estimated the SNR in the input video. We briefly discuss the denoising techniques used in Eulerian linear video magnification and their effect on this SNR. The Eulerian linear video magnification algorithm uses a spatial averaging filter (with a Gaussian kernel) and a temporal band pass filter to isolate and denoise the signal of interest [1]. Each of these filters attenuates the variance of additive white Gaussian noise. We assume that the magnification step that is applied after these filters has a negligible effect on the output SNR.

Temporal filter

Under the assumption that the noise in the video can be modeled as additive white Gaussian noise, applying an ideal temporal band pass filter reduces the noise power by a factor that is proportional to the width of the filter [75]. We assume for now that this improves the SNR by the same factor; we will address this assumption shortly. We approximate the effect of this filter on the SNR:

$$SNR'(\boldsymbol{x}) = 10\log_{10}\left(\frac{F_s}{2w}\right)\left(\frac{\hat{\sigma}_s^2(\boldsymbol{x})}{\hat{\sigma}_n^2(\boldsymbol{x})}\right)$$

where F_s is the sampling rate and w is the width of the pass band. We assume $F_s = 30$ Hz. In [1], a band pass filter of $w = 0.17$ Hz is used to visualize blood flow in the face. In our work, we assume $w = 0.3$ Hz, since the resting heart rate in healthy individuals has been observed to vary by 0.12 Hz or more [76]. This value is discussed in further

detail in section 3.2. Under these assumptions, the SNR after the application of the temporal filter becomes:

$$SNR'(\boldsymbol{x}) = 10 \log_{10} \left(50 \cdot \frac{\hat{\sigma}_s^2(\boldsymbol{x})}{\hat{\sigma}_n^2(\boldsymbol{x})} \right)$$

To check the validity of this assumption, we measured the effect of a 2nd-order Butterworth filter on simulated white Gaussian noise. We found that the filter reduced the noise power by a mean factor of 60. This indicates that our assumed noise power reduction factor of 50 is reasonable. We use the assumed factor of 50 to make a more conservative estimate of the output SNR.

Spatial averaging filter

While the Eulerian linear video magnification algorithm uses a Gaussian kernel in the spatial averaging step, we consider the box filter in our work. Our techniques can easily be adapted to spatial filters of other shapes.

We saw in section 2.4.2 that applying a box filter of size $m \times m$ modifies the variance of additive white Gaussian noise by a factor of $\frac{1}{m^2}$. This produces the output SNR:

$$SNR_{out}(\boldsymbol{x}) = 10 \log_{10} \left(50 \cdot m^2 \cdot \frac{\hat{\sigma}_s^2(\boldsymbol{x})}{\hat{\sigma}_n^2(\boldsymbol{x})} \right)$$

In section 2.4.2, we discussed the importance of choosing an appropriate value for m. While applying a larger filter attenuates more noise, too much spatial averaging will distort the signal of interest. The appropriate threshold for m depends on the spatial characteristics of the signal of interest as well as the resolution of the video. For the purposes of visualizing blood flow in the hand, we choose m to be equal to the width of the index finger (this is an arbitrary heuristic and should be investigated in future work). In our video, the width of the index finger is approximately 30px.

Using $m = 30$,

$$SNR_{out}(\boldsymbol{x}) = 10 \log_{10} \left(45000 \cdot \frac{\hat{\sigma}_s^2(\boldsymbol{x})}{\hat{\sigma}_n^2(\boldsymbol{x})} \right)$$

2.6.6 Output SNR

We predicted that applying the temporal and spatial filters described in section 2.6.5 produces an average output SNR of 33.9 dB in the hand (Fig. 2-17). On average, the signal amplitude is in the resulting video is predicted to be over 60 times larger than the noise standard deviation. It is likely that this estimate was affected by the motion in the hand, which we discussed in previous sections. The accuracy of this estimate also depends on the validity of our assumptions. We assumed that applying a temporal filter would attenuate the noise without affecting the signal power. However, if the signal has non-zero power outside of the pass band of the temporal filter, applying the filter attenuates the signal power as well as the noise power. It is also possible that using a spatial averaging filter with a width equal to that of the index finger distorts the signal of interest and decreases the output SNR.
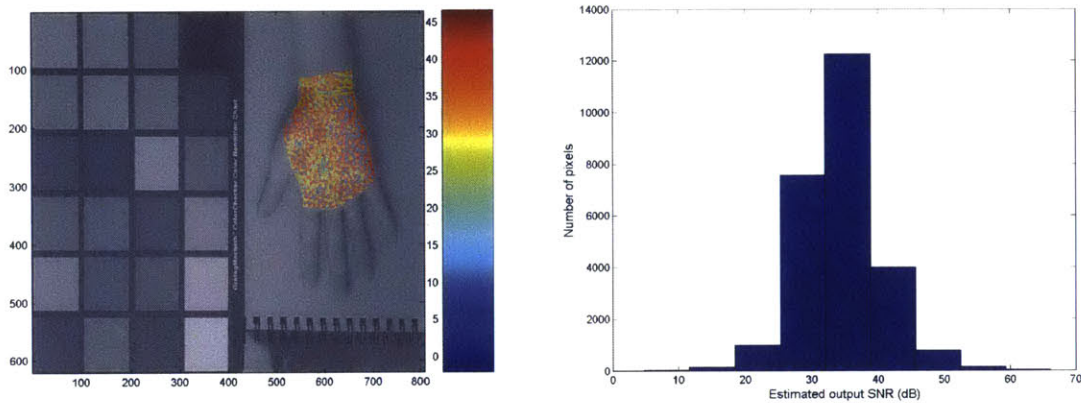
(a) Estimated output SNR (dB) at each pixel in the hand.

(b) Histogram of estimated output SNRs (dB) in the hand.

Figure 2-17: Using the temporal and spatial filters described earlier, we estimate that we can obtain high output SNRs in the hand. We predicted that these filters increase the SNR from an mean of -12.1 dB in the input video to a mean of 33.9 dB in the output video.

2.7 Summary

In this chapter, we discussed a model for the signal and noise terms in video magnification. We described an algorithm for estimating the amplitude of the signal of interest in the context of visualizing blood flow. We showed that the algorithm could be used to accurately estimate the signal amplitude in a video with a low SNR. We discussed how the parameters used in Eulerian linear video magnification affect the SNR of the resulting video. We applied our algorithm to a video of a hand, estimated the input SNR, and predicted the output SNR. While our methods are tailored to the task of visualizing blood flow, the techniques we presented can be adapted to other applications of video magnification.

Chapter 3

Visualizations of blood flow

When attempting to examine blood flow for the purpose of diagnosing or tracking peripheral arterial disease, our signal of interest is the color changes in the skin that that occur as a result of blood perfusion. In this section, we present some techniques for visualizing this signal. We evaluate the accuracy of these visualizations, and investigate whether they can be used to differentiate a healthy extremity from an extremity affected by an arterial occlusion.

Peripheral arterial disease can affect any of the arteries outside of the heart, including the carotid artery and arteries in the brain, upper extremities, kidney and lower extremities [7]. The lower extremity vessels are affected more commonly than the upper extremity vessels [18]. However, we focus our early investigations on the hands since they are easy to access, manipulate and film.

3.1 Vasculature of the hands

In this section, we briefly review the vasculature of the human hand. The human hand is supplied by two arteries: the ulnar artery (located on the same side of the hand as the pinky) and the radial artery (located on the same side of the hand as the thumb). As shown in Fig. 3-1, these arteries meet in the center of the hand to form the deep palmar arch and the superficial palmar arch [78]. The deep palmar arch is supplied mostly by the radial artery, and branches into the vessels that supply the

thumb and radial side of the index finger. The superficial palmar arch is supplied mostly by the ulnar artery, and in turn supplies all of the fingers but the thumb. Most of the arteries in the hand are located on the palmar side. The venous system of the hand is mostly located on the dorsal side [79].
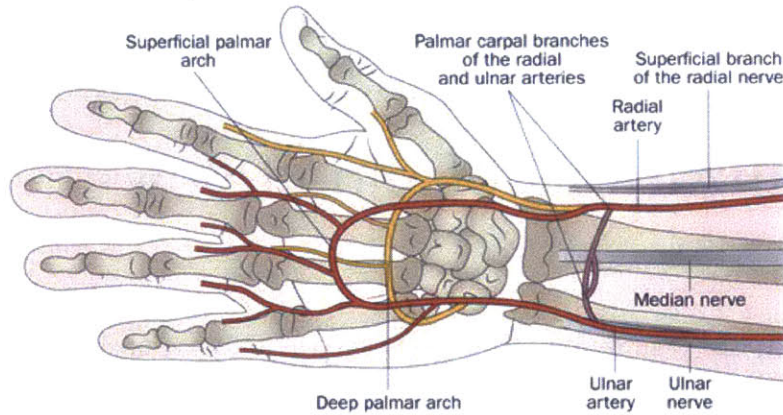
Figure 3-1: Vasculature of the hand, viewed from the palmar side. Figure reproduced from [80].

We found that applying linear video magnification to videos of the dorsal side of the hand produced no color changes at the heart rate. This is consistent with the vascular anatomy, as most of the arteries of the hand lie on the palmar side of the muscles and bones of the hand [81].

3.2   Color change as the signal of interest

In the application of visualizing blood flow, the signal of interest is the color change in the skin resulting from the underlying flow of blood. In this section, we make assumptions about the temporal and spatial characteristics of this signal. We hypothesize that these assumptions will help us isolate and visualize the signal of interest.

There is sufficient diagnostic power in a small band of frequencies near the heart rate

Blood pressure waveforms can have non-zero power in multiple frequency bands, as shown in Fig. 3-2. It is likely that the color change resulting

70

from the propagation of this pulse through small blood vessels has similar characterstics. However, we assume that the components of the signal at frequencies near the heart rate contain sufficient information to produce an accurate visualization of blood flow. In the frequency domain, we consider only the signal components in a band of width w_h centered at the heart rate f_h. This band is intended to describe the natural variations in resting heart rate in healthy subjects. The 95% confidence interval for the difference in resting heart rate varies from 7 beats/min (or 0.12 Hz) in 20-year-olds to 3 beats/min (or 0.05 Hz) in 75 year olds [76]. In our work, we allow for variations of up to 0.15 Hz around the heart rate, and assume $w_h = 2 \times 0.15$ Hz $= 0.3$ Hz.
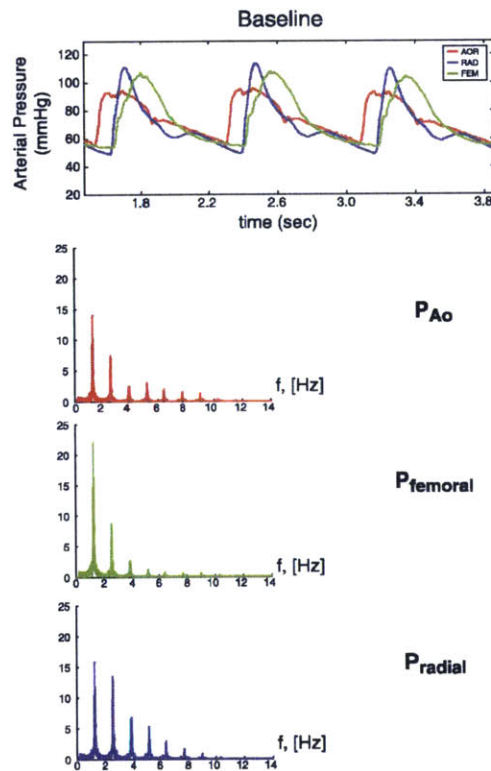
Figure 3-2: Arterial pressure in the aorta (P_{Ao}), femoral artery ($P_{femoral}$) and radial artery (P_{radial}) of a female pig, in the time domain and frequency domain. Figure adapted from [82].

The signal varies gradually across space

> We assume that each pixel exhibits a similar signal of interest to the pixels in a small neighborhood around it. This assumption allows us to apply spatial averaging filters up to a certain size to the video without significantly changing the signal amplitude at each pixel. In our work, we choose the maximum filter side length to be equal to the width of the index finger. Future work should investigate the validity of this assumption in healthy individuals as well as individuals with vascular abnormalities.

These assumptions allow us to determine some of the parameters used in video magnification and in other visualizations, which we will see in the following sections.

3.3 Application of Eulerian linear video magnification

In this section, we use Eulerian linear video magnification to visualize the blood flow in two human hands. We discuss whether the magnified videos accurately reflect blood flow. We also investigate whether the results can be used to differentiate a healthy hand from a hand affected by an arterial occlusion.

We produced a video of the hand of a healthy female subject, and the hand of a female subject who had an occlusion in the ulnar artery at the wrist, as verified by MRA. Each of the videos was taken in a room with consistent, bright lighting. In each video, the subject was positioned comfortably next to a flat surface that was approximately level with her chest. The subject was instructed to place her forearm flat on the surface and to keep it as still as possible for the duration of the video. In the case of the healthy subject, we placed a handheld continuous-wave Doppler ultrasound probe (which we will henceforth refer to as the Doppler probe) on the subject's radial artery to obtain an audio recording of the subject's pulse.

We processed the videos using a modified version of the Eulerian linear video magnification algorithm presented in [1]. In [1], all three color channels magnified.

Figure 3-3: Experiment setup with a healthy female subject. The Doppler probe was placed on the radial artery in the wrist.

Our algorithm only magnifies the green channel, as it was described by Verkruysse *et al.* to contain the strongest signal relating to heart activity [27].

3.3.1 Linear video magnification algorithm for visualizing blood flow in the hand

We processed each video in Matlab according to the following algorithm:

1. Manually designate a region of interest that includes the palm of the hand and the wrist. This region should exclude the background, the fingers (which are prone to unwanted motions), and the edges of the hand.

2. Examine the time series of all the pixels within the region of interest. Select a video segment that does not contain large changes in amplitude; these are likely caused by motion.

3. Examine the frequency spectrum of all the pixels for the selected video segment, and confirm that there is a clear peak within the range of 0.5 Hz to 2.0 Hz.

Designate f_h to be the frequency at which this peak occurs. If there is no peak, return to the previous step and select another video segment.

4. For the green channel of the selected video segment:

   (a) Apply a spatial blur using a Gaussian kernel. The standard deviation of the Gaussian, σ_f, should be approximately equal to half of the width of the index finger; the exact value can be determined experimentally.

   (b) Apply a temporal band-pass filter to exclude frequencies outside of $[f_h - \frac{w_h}{2}, f_h + \frac{w_h}{2}]$ where $w_h = 0.3$ Hz.

   (c) Magnify the intensity of each pixel by a constant factor, α.

5. Add the magnified green channel to the original green channel.

6. Recombine the color channels to create the magnified video.

3.3.2 Results

Healthy hand

The results for the healthy subject are shown in Fig. 3-4. In examining the frequency spectra of the pixels in the region of interest, we found a clear peak at 1.17 Hz, which we verified to be the heart rate using the audio signal from the Doppler device. Fig. 3-4a shows that the hand uniformly exhibits a green to red color change at the heart rate frequency; this is similar to the blood flow effect that has been observed in the face[1].

Fig. 3-4b indicates that the minima in the green channel intensity exhibit a fairly consistent time delay compared to the peaks in the audio signal. We investigate this time delay in the next section. This alignment suggests that this particular magnified video reflects the blood flow detected by the Doppler device. Due to the relatively narrow temporal pass-band that we applied to the video, the intensity in the magnified video is forced to vary at a rate that is close to 1.17 Hz. Furthermore, the beat-to-beat variations in heart rate that are visible in the audio signal (highlighted by the red box)
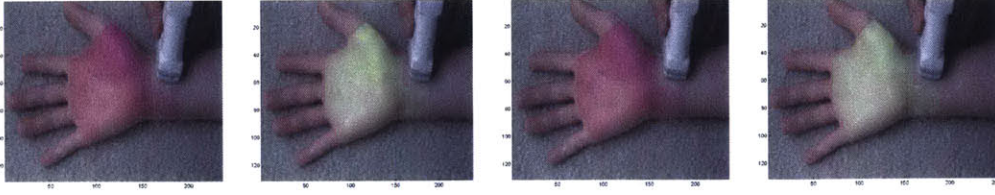
(a) Frames of a magnified video of a healthy hand, separated by time intervals of $\frac{1}{2f_h}$ seconds. The heart rate f_h was determined from the frequency spectra of the pixels in the region of interest to be 1.17 Hz.
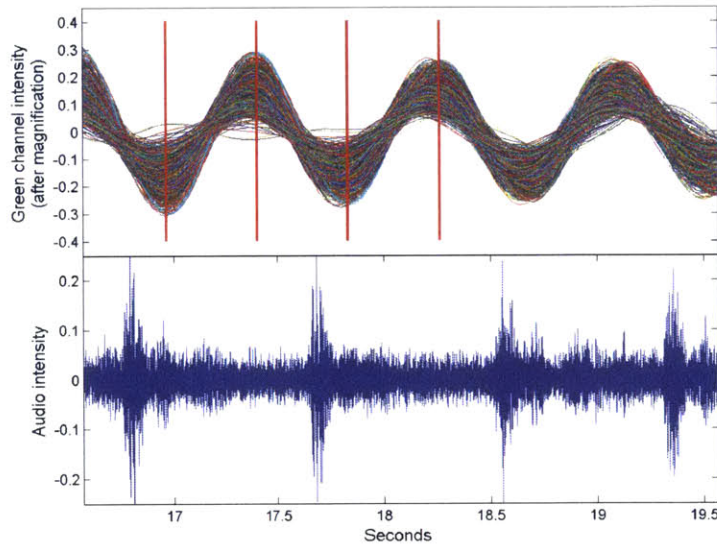
(b) (Top) Green channel intensity of the pixels in the region of interest of the magnified video, which produced at the end of step 4 in section 3.3.1). Each curve represents the intensity of a single pixel. The red lines represent the timing of the four frames presented in Fig. 3-4a. (Bottom) Audio signal produced by the Doppler device.

Figure 3-4: In the magnified video of a healthy hand, we see that the hand uniformly exhibits a green to red color change at the heart rate frequency.

are not apparent in the magnified green channel. Future work should investigate the performance of the video magnification algorithm on videos that contain large beat-to-beat variations in heart rate.

Hand with ulnar artery occlusion

The results for the subject with an occluded radial artery are shown in Fig. 3-6. We found the heart rate $f_h = 1.29$ Hz in a similar fashion to the previous video.

75
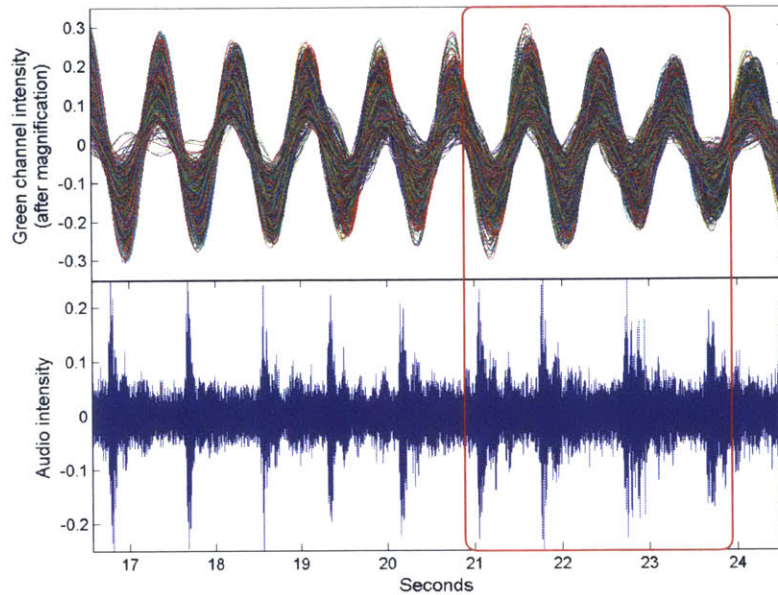
Figure 3-5: The intensity in the magnified green channel (each curve represents the signal at a single pixel) appears to align with the peaks in the audio produced by the Doppler device. In the time segment highlighted in red, however, the magnified intensity does not exhibit the same beat-to-beat variations in heart rate that are shown in the audio signal.

Unfortunately, we did not have access to a continuous-wave Doppler ultrasound device at the time the input video was taken. In the magnified video, the hand does not appear to exhibit a uniform color change. There appears to be a high concentration of color at the ulnar part of the palm near the wrist (shown in the blue box in 3-6a) that is not apparent in the healthy hand (3-4a). This discrepancy may be caused by the occlusion in the ulnar artery.

In this pilot study, we showed that the Eulerian linear video magnification technique produces magnified videos that appear to reflect the blood flow in a hand. These videos also show some differences between a healthy hand and a hand with an occluded ulnar artery. Further work is required to determine if these visual differences are caused by differences in blood flow patterns between the two subjects. In the next section, we explore an alternate visualization of blood flow.
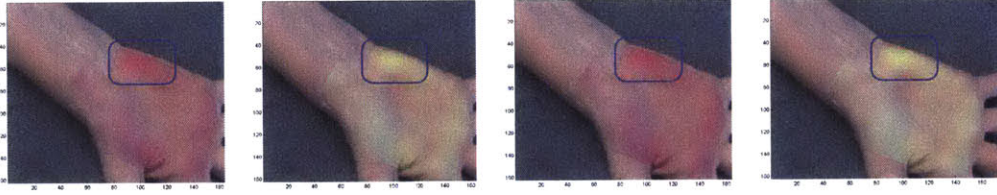
(a) Frames of a magnified video of a hand with an occluded radial artery, separated by time intervals of $\frac{0.5}{f_h}$ seconds. The heart rate f_h was determined from the frequency spectra of the pixels in the region of interest to be 1.29 Hz. The blue box indicates an area that appears to exhibit large color changes.
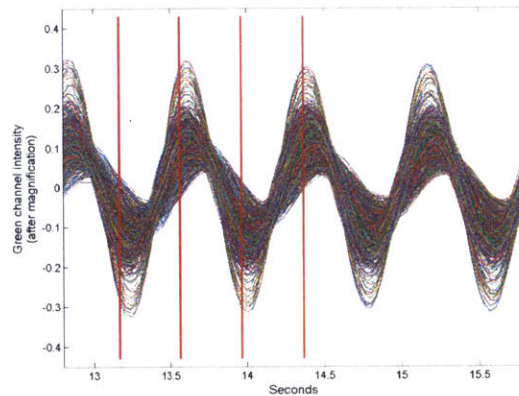
(b) Green channel intensity of the pixels in the region of interest of the magnified video. Each curve represents the intensity of a single pixel. The red lines represent the timing of the four frames presented in Fig. 3-4a.

Figure 3-6: In the magnified video of a hand with an occluded ulnar artery, we see that the hand does not exhibit uniform color changes at the heart rate frequency.

3.4   Visualization of the time delay of blood flow

The amount of time that it takes for blood to perfuse the skin may be indicative of a nearby arterial occlusion. In continuous-wave Doppler ultrasound recordings, the pulse transit time describes the time it takes for the pulse wave in a cardiac cycle to travel between two sites along an artery [83]. An increased pulse transit time can indicate the blockage of an artery [83, 84]. In this section, we use magnified videos to produce visualizations of the time delay between the signal of interest at different parts of the hand. We investigate the differences in the results produced for a healthy hand and a hand affected by an ulnar artery occlusion.

3.4.1 Cross-correlation for determining time delay

The cross-correlation method is the most common way of determining the time delay between two signals [85]. In [86], the cross-correlation of two sequences $y_1(t)$ and $y_2(t)$ is defined as:

$$R_{y_1 y_2}(\tau) = \mathrm{E}[y_1(t)y_2(t - \tau)]$$

For finite-length signals, the cross-correlation estimated for the observation interval T is defined as:

$$\hat{R}_{y_1 y_2}(\tau) = \frac{1}{T - \tau} \int_\tau^T y_1(t)y_2(t - \tau)dt$$

The value of τ that maximizes $\hat{R}_{y_1 y_2}$ is the estimate of the delay D between the signals. The input signals $y_1(t)$ and $y_2(x, t)$ may be filtered using H_1 and H_2 respectively prior to performing cross-correlation in order to improve the accuracy of the delay estimate [86]. This algorithm is summarized in Fig. 3-7.
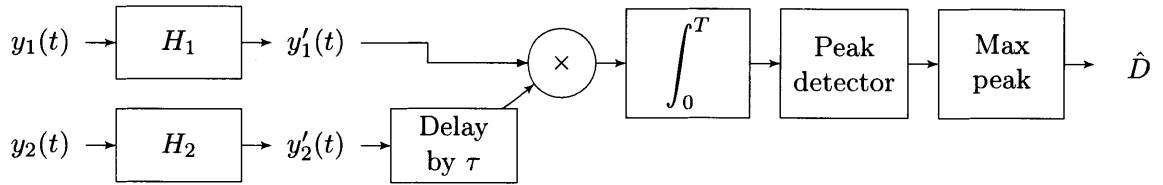
Figure 3-7: Cross-correlation algorithm for computing the time delay between two signals. The input signals $y_1(t)$ and $y_2(t)$ may be filtered to produce $y_1'(t)$ and $y_2'(t)$. Figure adapted from [86].

3.4.2 Cross-correlation in magnified videos

In the visualization of blood flow, our goal is to estimate the time delay at each pixel, $D(x)$. We use the cross-correlation algorithm described in Fig. 3-7 to compare the green channel intensity signal at each pixel x to some fiduciary signal $y_1(t)$.

We define the $y_1(t)$ to be the green channel intensity signal at some reliable pixel x_r, which we call the fiduciary point. We define "reliable" as having a clear and high-magnitude peak at f_h in the frequency domain. The location of this fiduciary point is determined by examining the height and sharpness of the frequency spectrum at

f_h for each pixel within the region of interest. We use $y_1'(t) = I_{mag,G}(\boldsymbol{x}_r, t)$, where $I_{mag,G}(\boldsymbol{x}, t)$ is the magnified intensity in the green channel produced by the Eulerian linear video magnification algorithm. This is equivalent to filtering $y_1(\boldsymbol{x}_r, t)$ using the spatial blur and temporal filter described in section 3.3.1.

For each $\boldsymbol{x}$ within the region of interest, we use $y_2'(\boldsymbol{x}, t) = I_{mag,G}(\boldsymbol{x}, t)$. We explore all time delays in the range $[-\frac{1}{2f_h}, \frac{1}{2f_h})$, that is, for all τ in $[-\frac{F_s}{2f_h}, \frac{F_s}{2f_h})$.

3.4.3 Results

For both subjects, we produced time delay estimates relative to a fiduciary point on the ulnar side of the hand, close to the wrist. The pixels in this area exhibited clear peaks at f_h for both subjects.

Fig. 3-8 shows noticeable differences between the two subjects. In the video of the healthy subject, the estimated time delays appear to be uniform throughout the palm (Fig. 3-8a). This is consistent with our earlier observation that in the magnified video, the healthy hand appears to uniformly change in color from green to red. In the video of the subject with an arterial occlusion, the time delays appear to be less homogeneous throughout the palm (Fig. 3-8c). Areas closer to the center of the hand have more negative delays compared to the edges of the hand.

Fig. 3-8d shows that the time delays for the healthy subject have a relatively narrow spread, and are centered approximately at 0. In contrast, the estimated time delays for the subject with an arterial occlusion (Fig. 3-8b) have a wider spread with two peaks. The time delays also appear to be shifted towards more negative values.

While there are noticeable differences between the videos of the two subjects, it is unclear if the differences are significant or contain diagnostic power. Further study is required to determine if these differences are caused by the arterial occlusion in the second subject.
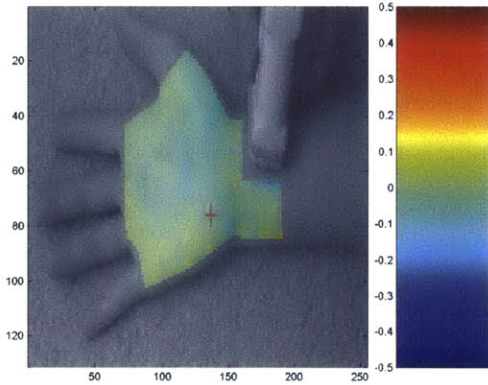
(a) Estimated time delays relative to fiduciary point (denoted by magenta '+') in the healthy hand, in multiples of $\frac{1}{f_h}$ where $f_h = 1.17$ Hz.
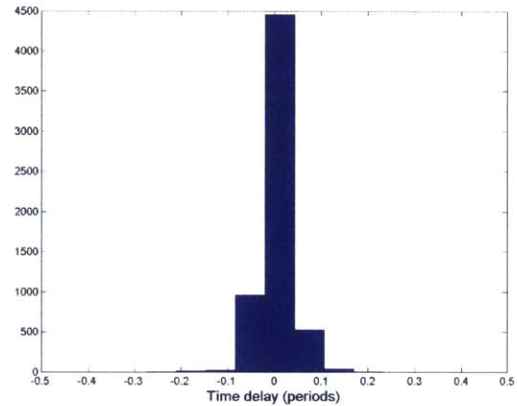
(b) Histogram of estimated time delays relative in the healthy hand.
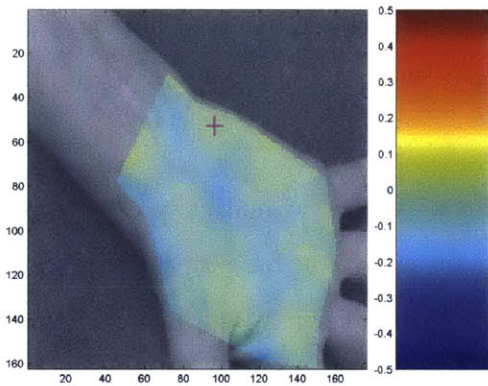
(c) Estimated time delays relative to fiduciary point (denoted by magenta '+') in hand with an occluded artery, in multiples of $\frac{1}{f_h}$ where $f_h = 1.29$ Hz.
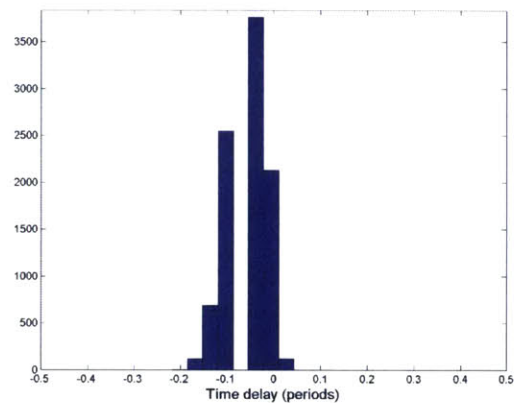
(d) Histogram of estimated time delays in the hand with an occluded artery.

Figure 3-8: The estimated time delay of the magnified green channel intensity signal in each pixel. These delays were computed over 250 video frames.

3.5 Summary

In this section, we presented several ways to examine blood flow in the hand. We produced visualizations from a video of a healthy hand, and a video of a hand with an occluded ulnar artery. The magnified video produced by the Eulerian linear video magnification algorithm appeared to agree with the audio signal produced

by a continuous-wave Doppler ultrasound device placed on the healthy subject's radial artery. The magnified videos showed several qualitative differences between the healthy hand and the hand affected by an arterial occlusion. We produced maps of the estimated time delay of the signal in each pixel in the hand in each video. These delay maps and their associated histograms indicated that the hand with an arterial occlusion had a higher spread of time delays. Future work should study the accuracy and reliability of these results.

Chapter 4

Conclusion

In this work, we addressed an important question about video magnification that had previously gone unanswered: how reliable are the visualizations produced by video magnification algorithms? Videos that are intended for video magnification typically have low SNRs. To our knowledge, no previous studies have measured the input or output SNR in video magnification.

We are particularly interested in producing accurate and reliable visualizations of blood flow in the human body. We presented a model to describe the signal and noise in the general context of video magnification, and then presented a more specific model to describe the signal and noise in the application of visualizing blood flow. Based on this model, we developed an algorithm for estimating the amplitude of the color change caused by blood perfusion in a video. We showed that this algorithm is able to accurately estimate the amplitude of a periodic color change in a simulated video with a low SNR. We applied the same algorithm to a real video of a healthy hand, and produced plausible estimates of the input SNR. Using these values, we predicted that the existing denoising methods used in video magnification were sufficient for producing a high output SNR. While our algorithm focused on the application of visualizing blood flow, it can be adapted to assess the SNR in other applications of video magnification. In future work, we would like to verify and improve the accuracy of our estimates and predictions.

We also demonstrated that video magnification can be used to produce visualiza-

tions that have noticeable differences between a healthy individual and an individual with a vascular abnormality. Such visualizations might be useful in a low-cost tool for evaluating vascular disease.

This work suggests that it is possible to obtain accurate visualizations of blood flow using video magnification. We believe that video magnification technology has the potential to make a meaningful impact in the evaluation of vascular diseases.

Bibliography

[1] Neal Wadhwa, Michael Rubinstein, Frédo Durand, and William T Freeman. Phase-based video motion processing. *ACM Transactions on Graphics (TOG)*, 32(4):80, 2013.

[2] Geoff Dougherty. *Digital image processing for medical applications*. Cambridge University Press, 2009.

[3] Maurice Gagnaire. *Broadband local loops for high-speed Internet access*. Artech House, 2002.

[4] Christophe Schlick. Quantization techniques for visualization of high dynamic range pictures. In *Photorealistic Rendering Techniques*, pages 7–20. Springer, 1995.

[5] TR Nelson and DH Pretorius. The doppler signal: where does it come from and what does it mean? *American Journal of Roentgenology*, 151(3):439–447, 1988.

[6] Jay Devore. *Probability and Statistics for Engineering and the Sciences*. Cengage Learning, 2015.

[7] Philip Lockwood and Jérôme Boudy. Experiments with a nonlinear spectral subtractor (nss), hidden markov models and the projection, for robust speech recognition in cars. *Speech communication*, 11(2):215–228, 1992.

[8] Boh Lim Sim, Yit Chow Tong, Joseph S Chang, and Chin Tuan Tan. A parametric formulation of the generalized spectral subtraction method. *Speech and Audio Processing, IEEE Transactions on*, 6(4):328–337, 1998.

[9] M Berouti, R Schwartz, and John Makhoul. Enhancement of speech corrupted by acoustic noise. In *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP'79.*, volume 4, pages 208–211. IEEE, 1979.

[10] Charles Poynton. *Digital video and HD: Algorithms and Interfaces*. Elsevier, 2012.

[11] International Electrotechnical Commission et al. Multimedia systems and equipment–colour measurement and management–part 2-1: Colour

management–default rgb colour space–srgb. Technical report, IEC 61966-2-1, 1999.

[12] Adobe Systems Incorporated. AdobeÂŏ rgb (1998) color image encoding. Technical report, Adobe Systems Incorporated, 2005.

[13] YL Lee, HC Kim, and HW Park. Blocking effect reduction of jpeg images by signal adaptive filtering. *Image Processing, IEEE Transactions on*, 7(2):229–234, 1998.

[14] IA O'Brien, Paul O'Hare, and RJ Corrall. Heart rate variability in healthy subjects: effect of age and the derivation of normal ranges for tests of autonomic function. *British heart journal*, 55(4):348–354, 1986.

[15] V Balakrishnan, Nash Borges, and Luke Parchment. Wavelet denoising and speech enhancement. *Department of Electrical and Computer Engineering, The Johns Hopkins University, Baltimore, Spring*, 2006.

[16] Sadaoki Furui. Speaker-independent isolated word recognition using dynamic features of speech spectrum. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 34(1):52–59, 1986.

[17] Patrenahalli M Narendra. A separable median filter for image noise smoothing. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (1):20–29, 1981.

[18] Feras Hatib, Jos RC Jansen, and Michael R Pinsky. Peripheral vascular decoupling in porcine endotoxic shock. *Journal of Applied Physiology*, 111(3):853–860, 2011.

[19] David B Jenkins. *Hollinshead's functional anatomy of the limbs and back*. Elsevier Health Sciences, 2008.

[20] Richard S Snell. *Clinical anatomy by systems*. Lippincott Williams & Wilkins, 2007.

[21] Abbas El Gamal and Helmy Eltoukhy. Cmos image sensors. *Circuits and Devices Magazine, IEEE*, 21(3):6–20, 2005.

[22] Rafael C Gonzalez and Richard E Woods. Digital image processing: second edition. pages 81–82, 2002.

[23] Rajeev Ramanath, Wesley E Snyder, Griff L Bilbro, and William A Sander. Demosaicking methods for bayer color arrays. *Journal of Electronic imaging*, 11(3):306–315, 2002.

[24] RG Gosling, G Dunbar, DH King, DL Newman, CD Side, JP Woodcock, DE Fitzgerald, JS Keates, and D MacMillan. The quantitative analysis of occlusive peripheral arterial disease by a non-intrusive ultrasonic technique. *Angiology*, 22(1):52–55, 1971.

[25] Charles Knapp and G Clifford Carter. The generalized correlation method for estimation of time delay. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 24(4):320–327, 1976.

[26] Maurizio Omologo and Piergiorgio Svaizer. Acoustic event localization using a crosspower-spectrum phase based technique. In *Acoustics, Speech, and Signal Processing, 1994. ICASSP-94., 1994 IEEE International Conference on*, volume 2, pages II–273. IEEE, 1994.

[27] JP Woodcock, RG Gosling, and DE Fitzgerald. A new non-invasive technique for assessment of superficial femoral artery obstruction. *British Journal of Surgery*, 59(3):226–231, 1972.

[28] Ce Liu, Antonio Torralba, William T Freeman, Frédo Durand, and Edward H Adelson. Motion magnification. *ACM Transactions on Graphics (TOG)*, 24(3):519–526, 2005.

[29] Jue Wang, Steven M Drucker, Maneesh Agrawala, and Michael F Cohen. The cartoon animation filter. In *ACM Transactions on Graphics (TOG)*, volume 25, pages 1169–1173. ACM, 2006.

[30] Justin G Chen, Neal Wadhwa, Young-Jin Cha, Frédo Durand, William T Freeman, and Oral Buyukozturk. Structural modal identification through high speed camera video: Motion magnification. In *Topics in Modal Analysis I, Volume 7*, pages 191–197. Springer, 2014.

[31] Neal Wadhwa, Michael Rubinstein, Fredo Durand, and William T Freeman. Riesz pyramids for fast phase-based video magnification. In *Computational Photography (ICCP), 2014 IEEE International Conference on*, pages 1–10. IEEE, 2014.

[32] Curt Diehm, Jens Rainer Allenberg, David Pittrow, Matthias Mahn, Gerhart Tepohl, Roman L Haberl, Harald Darius, Ina Burghaus, Hans Joachim Trampisch, et al. Mortality and vascular morbidity in older adults with asymptomatic versus symptomatic peripheral artery disease. *Circulation*, 120(21):2053–2061, 2009.

[33] Salvatore Novo. Classification, epidemiology, risk factors, and natural history of peripheral arterial disease. *Diabetes, Obesity and Metabolism*, 4(s2):S1–S6, 2002.

[34] Virginia A Moyer. Screening for peripheral artery disease and cardiovascular disease risk assessment with the ankle–brachial index in adults: Us preventive services task force recommendation statement. *Annals of internal medicine*, 159(5):342–348, 2013.

[35] Jeroen G Lijmer, Maria GM Hunink, Jan JAM van den Dungen, Jan Loonstra, and Andries J Smit. Roc analysis of noninvasive tests for peripheral arterial disease. *Ultrasound in medicine & biology*, 22(4):391–398, 1996.

[36] Rody Ouwendijk, Marianne de Vries, Theo Stijnen, Peter MT Pattynama, Marc RHM van Sambeek, Jaap Buth, Alexander V Tielbeek, Daan A van der Vliet, Leo J SchutzeKool, Peter JEHM Kitslaar, et al. Multicenter randomized controlled trial of the costs and effects of noninvasive diagnostic imaging in patients with peripheral arterial disease: the dipad trial. *American Journal of Roentgenology*, 190(5):1349–1357, 2008.

[37] Michael H Criqui, Mark J Alberts, F Gerald R Fowkes, Alan T Hirsch, Patrick T O'Gara, Jeffrey W Olin, et al. Atherosclerotic peripheral vascular disease symposium ii screening for atherosclerotic vascular diseases:: Should nationwide programs be instituted? *Circulation*, 118(25):2830–2836, 2008.

[38] Alan T Hirsch, Ziv J Haskal, Norman R Hertzer, Curtis W Bakal, Mark A Creager, Jonathan L Halperin, Loren F Hiratzka, WR Murphy, Jeffrey W Olin, Jules B Puschett, et al. Acc/aha 2005 practice guidelines for the management of patients with peripheral arterial disease (lower extremity, renal, mesenteric, and abdominal aortic): a collaborative report from the american association for vascular surgery/society for vascular surgery, society for cardiovascular angiography and interventions, society for vascular medicine and biology, society of interventional radiology, and the acc/aha task force on practice guidelines (writing committee to develop guidelines for the management of patients with peripheral arterial disease): endorsed by the american association of cardiovascular and pulmonary rehabilitation; national heart, lung, and blood institute; society for vascular nursing; transatlantic inter-society consensus; and vascular disease foundation. *Circulation*, 113(11):e463–654, 2006.

[39] Donald Lloyd-Jones, Robert J Adams, Todd M Brown, Mercedes Carnethon, Shifan Dai, Giovanni De Simone, T Bruce Ferguson, Earl Ford, Karen Furie, Cathleen Gillespie, et al. Executive summary: heart disease and stroke statistics–2010 update: a report from the american heart association. *Circulation*, 121(7):948, 2010.

[40] F Gerald R Fowkes, Diana Rudan, Igor Rudan, Victor Aboyans, Julie O Denenberg, Mary M McDermott, Paul E Norman, Uchechukwe KA Sampson, Linda J Williams, George A Mensah, et al. Comparison of global estimates of prevalence and risk factors for peripheral artery disease in 2000 and 2010: a systematic review and analysis. *The Lancet*, 382(9901):1329–1340, 2013.

[41] Elizabeth Selvin and Thomas P Erlinger. Prevalence of and risk factors for peripheral arterial disease in the united states results from the national health and nutrition examination survey, 1999–2000. *Circulation*, 110(6):738–743, 2004.

[42] Hao-Yu Wu, Michael Rubinstein, Eugene Shih, John V Guttag, Frédo Durand, and William T Freeman. Eulerian video magnification for revealing subtle changes in the world. *ACM Trans. Graph.*, 31(4):65, 2012.

[43] Ce Liu, William T Freeman, Richard Szeliski, and Sing Bing Kang. Noise estimation from a single image. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, volume 1, pages 901–908. IEEE, 2006.

[44] Aishy Amer and Hartmut Schroder. A new video noise reduction algorithm using spatial subbands. In *Electronics, Circuits, and Systems, 1996. ICECS'96., Proceedings of the Third IEEE International Conference on*, volume 1, pages 45–48. IEEE, 1996.

[45] Soeren I. Olsen. Estimation of noise in images: an evaluation. *CVGIP: Graphical Models and Image Processing*, 55(4):319–323, 1993.

[46] Aishy Amer and Eric Dubois. Fast and reliable structure-oriented video noise estimation. *Circuits and Systems for Video Technology, IEEE Transactions on*, 15(1):113–118, 2005.

[47] Kostadin Dabov, Alessandro Foi, and Karen Egiazarian. Video denoising by sparse 3d transform-domain collaborative filtering. In *Proc. 15th European Signal Processing Conference*, volume 1, page 7, 2007.

[48] James C Brailean, Richard P Kleihorst, Serafim Efstratiadis, Aggelos K Katsaggelos, and Reginald L Lagendijk. Noise reduction filters for dynamic image sequences: a review. *Proceedings of the IEEE*, 83(9):1272–1292, 1995.

[49] Michihiro Kobayashi, Takahiro Okabe, and Yoichi Sato. Detecting forgery from static-scene video based on inconsistency in noise level functions. *Information Forensics and Security, IEEE Transactions on*, 5(4):883–892, 2010.

[50] HG Hirsch and C Ehrlicher. Noise estimation techniques for robust speech recognition. In *Acoustics, Speech, and Signal Processing, 1995. ICASSP-95., 1995 International Conference on*, volume 1, pages 153–156. IEEE, 1995.

[51] Santiago Aja-Fernández, Gonzalo Vegas-Sánchez-Ferrero, Marcos Martín-Fernández, and Carlos Alberola-López. Automatic noise estimation in images using local statistics. additive and multiplicative cases. *Image and Vision Computing*, 27(6):756–770, 2009.

[52] Antoni Buades, Bartomeu Coll, and J-M Morel. A non-local algorithm for image denoising. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 2, pages 60–65. IEEE, 2005.

[53] Ljubomir Jovanov, Aleksandra Pizurica, Stefan Schulte, Peter Schelkens, Adrian Munteanu, Etienne Kerre, and Wilfried Philips. Combined wavelet-domain and motion-compensated video denoising based on video codec motion estimation methods. *Circuits and Systems for Video Technology, IEEE Transactions on*, 19(3):417–421, 2009.

[54] Kenji Irie, Alan E McKinnon, Keith Unsworth, and Ian M Woodhead. A technique for evaluation of ccd video-camera noise. *Circuits and Systems for Video Technology, IEEE Transactions on*, 18(2):280–284, 2008.

[55] Jan Lukas, Jessica Fridrich, and Miroslav Goljan. Digital camera identification from sensor pattern noise. *Information Forensics and Security, IEEE Transactions on*, 1(2):205–214, 2006.

[56] Ronan O'Rahilly and Fabiola Müller. *Basic human anatomy: a regional study of human structure*. WB Saunders Co, 1983.

[57] SC Ling and HB Atabek. A nonlinear analysis of pulsatile flow in arteries. *Journal of Fluid Mechanics*, 55(03):493–511, 1972.

[58] J Lindert, J Werner, M Redlin, H Kuppe, H Habazettl, and AR Pries. Ops imaging of human microcirculation: a short technical report. *Journal of vascular research*, 39(4):368–372, 2001.

[59] Jerome Aroesty and Joseph Francis Gross. The mathematics of pulsatile flow in small vessels i. casson theory. *Microvascular research*, 4(1):1–12, 1972.

[60] Keshen R Mathura, Karlijn C Vollebregt, Kees Boer, Jurgen C De Graaff, Dirk T Ubbink, and Can Ince. Comparison of ops imaging and conventional capillary microscopy to study the human microcirculation. *Journal of Applied Physiology*, 91(1):74–78, 2001.

[61] Mary P Wiedeman. Dimensions of blood vessels from distributing artery to collecting vein. *Circulation research*, 12(4):375–378, 1963.

[62] Wilmer Nichols, Michael O'Rourke, and Charalambos Vlachopoulos. *McDonald's blood flow in arteries: theoretical, experimental and clinical principles*. CRC Press, 2011.

[63] Collin Stultz. Using video magnification to visualizing blood flow. In-person discussion, April 2015.

[64] Luca Remonda, Pascal Senn, Alain Barth, Marcel Arnold, Karl-Olof Lövblad, and Gerhard Schroth. Contrast-enhanced 3d mr angiography of the carotid artery: comparison with conventional digital subtraction angiography. *American journal of neuroradiology*, 23(2):213–219, 2002.

[65] Stefan O Schoenberg, Johannes Rieger, Christof H Weber, Henrik J Michaely, Tobias Waggershauser, Carina Ittrich, Olaf Dietrich, and Maximilian F Reiser. High-spatial-resolution mr angiography of renal arteries with integrated parallel acquisitions: Comparison with digital subtraction angiography and us 1. *Radiology*, 235(2):687–698, 2005.

[66] Michael J Shea. Blood vessels. Merck Manual: Consumer Version, August 2013.

[67] Perry Carter. The cardiovascular system: Blood vessels. Midlands Technical College Bio 211 Lecture Notes, unknown.

[68] Lauralee Sherwood. *Fundamentals of human physiology*. Cengage Learning, 2011.

[69] Stanley S Franklin, Martin G Larson, Shehzad A Khan, Nathan D Wong, Eric P Leip, William B Kannel, and Daniel Levy. Does the relation of blood pressure to coronary heart disease risk change with aging? the framingham heart study. *Circulation*, 103(9):1245–1249, 2001.

[70] William B Kannel, Tavia Gordon, and Melvin J Schwartz. Systolic versus diastolic blood pressure and risk of coronary heart disease: the framingham study. *The American journal of cardiology*, 27(4):335–346, 1971.

[71] Matteo Maggioni, Giacomo Boracchi, Alessandro Foi, and Karen Egiazarian. Video denoising, deblocking, and enhancement through separable 4-d non-local spatiotemporal transforms. *Image Processing, IEEE Transactions on*, 21(9):3952–3966, 2012.

[72] Matan Protter and Michael Elad. Image sequence denoising via sparse and redundant representations. *Image Processing, IEEE Transactions on*, 18(1):27–35, 2009.

[73] Keigo Hirakawa and Thomas W Parks. Image denoising for signal-dependent noise. In *ICASSP (2)*, pages 29–32, 2005.

[74] JS Lee and K Hoppel. Noise modeling and estimation of remotely-sensed images. In *Geoscience and Remote Sensing Symposium, 1989. IGARSS'89. 12th Canadian Symposium on Remote Sensing., 1989 International*, volume 2, pages 1005–1008. IEEE, 1989.

[75] Guha Balakrishnan, Fredo Durand, and John Guttag. Detecting pulse from head motions in video. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 3430–3437. IEEE, 2013.

[76] Sokwoo Rhee, Boo-Ho Yang, and Haruhiko Harry Asada. Artifact-resistant power-efficient design of finger-ring plethysmographic sensors. *Biomedical Engineering, IEEE Transactions on*, 48(7):795–805, 2001.

[77] Sunil Kamath and Philipos Loizou. A multi-band spectral subtraction method for enhancing speech corrupted by colored noise. In *IEEE international conference on acoustics speech and signal processing*, volume 4, pages 4164–4164. Citeseer, 2002.

[78] Antoni Buades, Bartomeu Coll, and Jean-Michel Morel. Image denoising methods. a new nonlocal principle. *SIAM review*, 52(1):113–147, 2010.

[79] Mukesh C Motwani, Mukesh C Gadiya, Rakhi C Motwani, and Frederick C Harris. Survey of image denoising techniques. In *Proceedings of GSPX*, pages 27–30, 2004.

[80] Pietro Perona and Jitendra Malik. Scale-space and edge detection using anisotropic diffusion. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 12(7):629–639, 1990.

[81] Ce Liu and William T Freeman. A high-quality video denoising algorithm based on reliable motion estimation. In *Computer Vision–ECCV 2010*, pages 706–719. Springer, 2010.

[82] Antoni Buades, Bartomeu Coll, and Jean-Michel Morel. Nonlocal image and movie denoising. *International journal of computer vision*, 76(2):123–139, 2008.

[83] Alastair JJ Wood and William R Hiatt. Medical treatment of peripheral arterial disease and claudication. *New England Journal of Medicine*, 344(21):1608–1621, 2001.

[84] Erica Mitchell Emile R Mohler III. Noninvasive diagnosis of arterial disease. IN: UpToDate, July 2012.

[85] Rebecca Basset. Finger and thumb anatomy. IN: UpToDate, February 2015.

[86] Maciej Dryjski Linda Harris. Epidemiology, risk factors, and natural history of peripheral artery disease. IN: UpToDate, January 2015.

[87] Rodney A Hayward. Screening for lower extremity peripheral artery disease. IN: UpToDate, May 2014.

[88] Michael A. Golden David G. Neschis. Clinical features and diagnosis of lower extremity peripheral artery disease. IN: UpToDate, June 2014. This is a full MISC entry.

[89] Amy B. Reed. Popliteal artery aneurysm. IN: UpToDate, November 2013. This is a full MISC entry.

[90] FGR Fowkes, E Housley, EHH Cawood, CCA Macintyre, CV Ruckley, and RJ Prescott. Edinburgh artery study: prevalence of asymptomatic and symptomatic peripheral arterial disease in the general population. *International journal of epidemiology*, 20(2):384–392, 1991.

[91] Jennifer S Lin, Carin M Olson, Eric S Johnson, Caitlyn A Senger, Clara B Soh, and Evelyn P Whitlock. The ankle brachial index for peripheral artery disease screening and cardiovascular disease prediction in asymptomatic adults: A systematic evidence review for the us preventive services task force. 2013.

[92] Francisco J Serrano Hernando and Antonio Martín Conejero. Peripheral artery disease: pathophysiology, diagnosis and treatment. *Revista Española de Cardiología*, 60(09):969–982, 2007.

[93] Lars Norgren, William R Hiatt, JA al Dormandy, Mark R Nehler, Kenneth A Harris, F Gerry R Fowkes, Tasc II Working Group, et al. Inter-society consensus for the management of peripheral arterial disease (tasc ii). *European Journal of Vascular and Endovascular Surgery*, 33(1):S1–S75, 2007.

[94] Bart S Ferket, Sandra Spronk, Ersen B Colkesen, and MG Myriam Hunink. Systematic review of guidelines on peripheral artery disease screening. *The American journal of medicine*, 125(2):198–208, 2012.

[95] Michael A Golden David G Neschis. Treatment of chronic lower extremity critical limb ischemia. IN: UpToDate, June 2014.

[96] Joshua A Beckman, Michael R Jaff, and Mark A Creager. The united states preventive services task force recommendation statement on screening for peripheral arterial disease more harm than benefit? *Circulation*, 114(8):861–866, 2006.

[97] Ros Collins, G Cranny, J Burch, R Aguiar-Ibanez, D Craig, K Wright, E Berry, M Gough, J Kleijnen, and M Westwood. A systematic review of duplex ultrasound, magnetic resonance angiography and computed tomography angiography for the diagnosis and assessment of symptomatic, lower limb peripheral arterial disease. *Health Technology Assessment*, 2007.

[98] Robert A Byrne, Salvatore Cassese, Maryam Linhardt, and Adnan Kastrati. Vascular access and closure in coronary angiography and percutaneous intervention. *Nature Reviews Cardiology*, 10(1):27–40, 2013.

[99] US Preventive Services Task Force et al. Screening for peripheral arterial disease: recommendation statement. *American family physician*, 73(3):497, 2006.

[100] José M Bioucas-Dias and Mário AT Figueiredo. Multiplicative noise removal using variable splitting and constrained optimization. *Image Processing, IEEE Transactions on*, 19(7):1720–1730, 2010.

[101] Glenn E Healey and Raghava Kondepudy. Radiometric ccd camera calibration and noise estimation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 16(3):267–276, 1994.

[102] Chris D Constantinides, Ergin Atalar, and Elliot R McVeigh. Signal-to-noise measurements in magnitude images from nmr phased arrays. *Magnetic Resonance in Medicine*, 38(5):852–857, 1997.

[103] MJ Firbank, A Coulthard, RM Harrison, and ED Williams. A comparison of two methods for measuring the signal to noise ratio on mr images. *Physics in medicine and biology*, 44(12):N261, 1999.

[104] Xinqiao Liu and Abbas El Gamal. Photocurrent estimation from multiple nondestructive samples in cmos image sensor. In *Photonics West 2001-Electronic*

Imaging, pages 450–458. International Society for Optics and Photonics, 2001. CMOS noise is also mostly additive.

[105] Hui Tian, Boyd Fowler, and Abbas E Gamal. Analysis of temporal noise in cmos photodiode active pixel sensor. *Solid-State Circuits, IEEE Journal of*, 36(1):92–101, 2001.

[106] Philipos C Loizou. *Speech enhancement: theory and practice*. CRC press, 2013.

[107] Israel Cohen. Relaxed statistical model for speech enhancement and a priori snr estimation. *Speech and Audio Processing, IEEE Transactions on*, 13(5):870–881, 2005.

[108] John G Proakis. *Intersymbol Interference in Digital Communication Systems*. Wiley Online Library, 2001.

[109] Steven Boll. Suppression of acoustic noise in speech using spectral subtraction. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 27(2):113–120, 1979.

[110] Ming-Zher Poh, Daniel J McDuff, and Rosalind W Picard. Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. *Optics express*, 18(10):10762–10774, 2010.

[111] John Allen. Photoplethysmography and its application in clinical physiological measurement. *Physiological measurement*, 28(3):R1, 2007.

[112] Wim Verkruysse, Lars O Svaasand, and J Stuart Nelson. Remote plethysmographic imaging using ambient light. *Optics express*, 16(26):21434–21445, 2008.

[113] Michal Tendera, Victor Aboyans, Marie-Louise Bartelink, Iris Baumgartner, Denis Clément, Jean-Philippe Collet, Alberto Cremonesi, Marco De Carlo, Raimund Erbel, F Gerry R Fowkes, et al. Esc guidelines on the diagnosis and treatment of peripheral artery diseases document covering atherosclerotic disease of extracranial carotid and vertebral, mesenteric, renal, upper and lower extremity arteriesthe task force on the diagnosis and treatment of peripheral artery diseases of the european society of cardiology (esc). *European heart journal*, 32(22):2851–2906, 2011.

[114] Peter J Burt and Edward H Adelson. The laplacian pyramid as a compact image code. *Communications, IEEE Transactions on*, 31(4):532–540, 1983.