MOTION MAGNIFICATION

submitted in partial fulfillment of the Degree of

BACHELOR OF TECHNOLOGY

IN

ELECTRONICS AND COMMUNICATION ENGINEERING

BY

Yashika Parmar (131068)
Sumeet Singh Chahal (131069)
Puvail Bansal (131070)

UNDER THE GUIDANCE OF
Dr. Neeru Sharma

Declaration

We hereby declare that the work reported in the B-Tech thesis entitled **"MOTION MAGNIFICATION"** submitted at **Jaypee University of Information Technology, Waknaghat, India**, is an authentic record of my work carried out under the supervision of Dr. Neeru Sharma.
We have not submitted this work elsewhere for any other degree or diploma.

Yashika Parmar

Sumeet Chahal

Puvail Bansal

Department of Electronics and Communication Engineering
Jaypee University of Information Technology, Waknaghat, India
Date – 20th April, 2017

Supervisor's Certificate

This is to certify that the work reported in the B-Tech. thesis entitled **"MOTION MAGNIFICATION"** submitted by Yashika Parmar, Sumeet Singh Chahal and Puvail Bansal at Jaypee University of Information Technology, Waknaghat, India, is a bona fide record of their original work carried out under my supervision.

This work has not been submitted elsewhere for any other degree or diploma.

(Signature)
Dr. Neeru Sharma

Date -

Acknowledgement

We would like to show our gratitude to everyone who helped us make this project successful, our teachers, support staff, fellow students. Also, special mention of the following must be made.

First of all, we would like to thank our project guide Dr. Neeru Sharma who helped us throughout this project and supported us. We are very thankful to her for giving her valuable time and attention to our project and helping us complete our project in time.

We are also thankful to our H.O.D. Dr. Sunil Bhooshan for giving us an opportunity to work on this project. Finally, we would like to thank all the members of staff who helped us and the lab faculty for allowing us to make use of lab equipment.

Yashika Parmar

Sumeet Chahal

Puvail Bansal

List of Figures

List of Abbreviations and Acronyms

- R&D - Research and Development
- SNR - Signal to Noise Ratio
- IIR - Infinite Impulse Response
- MATLAB - Mathematical Laboratory
- ISO - Measurement of Camera's sensititvity
- GUI - Graphical User Interface
- JPEG - Joint Photographic Experts Group
- MPEG - Moving Picture Experts Group
- DVMAG - Dynamic Video Motion Magnification
- PSNR - Peak Signal to Noise Ratio

Abstract

Everyday, several things around us go unnoticed as they cannot be detected by our eyes. In this thesis, we focus on a technique, video magnification which allows us to see these small or subtle movements of an object or even a human being, which are impossible to see through our naked eyes. This allows us to see everything around us with a new perspective. The method used here is Eulerian Video Magnification, which takes video of an object as the input, applies the spatial and temporal processing to it and then gives the output as the magnified version of the original input.

For the motion magnification to happen, it is very important that we measure visual motions with accuracy and also have a pair or more of pixels together that are to be modified. The final output after applying motion magnification reveals small, subtle motions in an amplified way that were impossible to detect in the original sequence. This method also requires the motions to be very small, which most of the time is not the case. Also, the motion of any layer can be magnified by a user specified value. Using this method, we will be able to visualize the amplified motion of the pulse rate of a hand or the flow of blood from heart to our face which changes our colour of the face. This method also runs in real time that shows the phenomenon, which is done by selecting temporal frequencies by the user and has a variety of applications in many fields.

Contents

Chapter 1

Introduction

1.1 Introduction

Visual movement for the most part happens at various amplitudes, and over various sizes of transient furthermore, spatial recurrence. Little movements, in actuality, are hard to watch, yet at the same time they may uncover imperative data about the world: little disfigurements of structures, moment alterations in a harmony procedure, or some little developments of a framework in light of some compelling capacity. We need a machine which will uncover and clear those movements, same as that of a magnifying instrument that uncovers little and undetectable structures. This procedure is called Motion Magnification, which acts like a magnifying instrument in video arrangements for movement. The calculation's principle idea is to investigate the movements introduced in an informative video arrangement, along these lines enabling the client to indicate a gathering of pixels that are to be taken a shot at, and to set how much their movements must be amplified. Ordinarily, little movements are made to be enhanced and expansive movements are left unaltered. The last phase of movement amplification is to bring about the grouping with the coveted movements to be amplified by the predefined sum. While movement amplification is thoughtfully straightforward, performing it without antiquities that are questionable requires significant tender loving care. We acquaint a few strategies with examine movement heartily, confirming evaluated movements and in addition their territories of support. The choice of movements to be increased is made basic and instinctive by a programmed gathering process, where a pre-determined number of pixel bunches are found, in light of their closeness in position, force, and movement attributes. To accomplish this, a measure of partiality is actualized. It's leeway is that it bunches focuses in light of their directions over a timeframe and not only the similitudes of their immediate speeds. The client is advised to indicate which bunch's movements ought to be increased and furthermore by how much. Gaps that are uncovered by opened up movements are generally filled utilizing surface combination techniques. Motion magnification has many potential applications that ranges from engineering diagnosis or instruction to comedic amplification of ordinary expressions.

1.2 Salient Features of Motion Magnification

Motion Magnification is a very interesting topic and has several unique Features. Some of these are

- Capable of magnifying seemingly invisible motions around us
- Can be used to detect motion easily, aiding surveillance.
- Has several applications in the field of Medicine and defense.
- Has a lot of raw untapped potential as very less work has been done on it.
- A very interesting R&D topic.

1.3 Objectives

A standard video can have small amplitude and hidden signals which cannot be observed with the naked eye because of limited sensitivity. These hidden signals variation can be very informative. This can be exaggerated to extract hidden information.
Previously, motion detection techniques have been developed to amplify color with low amplitude such as human skin color varies due to blood cir- culation which is invisible to the naked eye. By successful implementation of 9these tools we can reveal small or invisible signals in a video.
Some efforts have been previously made to amplify and unveil subtle motions that are invisible otherwise, using a cartoon animated filter to exaggerate motion. However, they use the Lagrangian approach, which is expensive in terms of computation and difficult to make it artifact free on boundaries for complicated motions.
A motion segmentation can be used to produce good quality synthesis but

this increases the complexity of an algorithm. The Eulerian prospective is to use properties of a voxel of fluids. Same differential approximation then forms a basis of optical flow algorithm. Temporal processes have been previously used to extract invisible signals. Eulerian motion magnification has the potential of being used as a monitoring and diagnostic application in medicine. It observes the hidden motions which are otherwise invisible otherwise to the naked eye.

Objectives of this thesis work are:

- To produce better results of motion magnification.
- To reduce noise and produce better quality output video.
- To magnify video using Gaussian and Steerable pyramid decompositions.
- To preserve video smoothness.
- To analyze and compare output videos.

1.4 Thesis Organization

Our Thesis is organized as follows:

- Chapter 1: This chapter introduces us to the basic concept of motion magnification. It also covers the salient features and objectives of this thesis.
- Chapter 2: In this chapter, we discuss the references used.
- Chapter 3: This chapter outlines the techniques used by us and analyses these techniques.
- Chapter 4: Here, we describe the steps used in motion magnification in detail.
- Chapter 5: This chapter concludes the thesis and verifies the results.

Chapter 2

Literature Review

Usually, for motion magnification, we require very subtle motion. But such a video may not be available in all cases. This may be due to background movement, noise or several other factors. In this paper, the author firstly talks about adding video stabilization to remove unwanted handshake and camera motion, but discards this approach as any error in video stabilization will also be magnified. DVMAG (Dynamic Video Motion Magnification) is a technique in which users manually select a region of interest to magnify which helps in motion magnification of videos with a wide range of motions[1]. In[2], the author takes a standard video sequence as input, and applies spatial decomposition, followed by temporal filtering to the frames. He then amplifies the resulting image to reveal hidden information. Using this technique, the author is able to visualize the flow of blood as it fills the face and also to amplify and reveal small motions. It can also run in real time to show phenomena occurring at temporal frequencies selected by the user. By analyzing the relation between motion and phase in steerable pyramids we show that by increasing the phase variations by a mul- tiplicative factor we can amplify subtle motions. We then use this relation to analyze the limits of our method, which are set by the spatial support of the steerable basis functions. To amplify motions further, we extend the complex steerable pyramid to sub-octavebandwidth pyramids, comprised of filters with larger spatial support in the primal domain. While our new image representation is over-complete by a larger factor, it supports larger amplification of motions at all spatial frequencies, leading to fewer artifacts[3]. A comparision between lagrangian based tracking of points over time and eulerian methods towards magnifying subtle video changes were proposed by decomposing the video frames spatially through bandpass filtering, and then temporally filtering the signal to find the information to be magnified[4]. The author then proves why eulerian methods are better due to their tendency to avoid amplification of unwanted objects in the background. Motion Magnification has also been sucessfully used for detecting pulse from color variations in the face of a subject. This has proven to be at par with hospital grade equipment and can sucessfully calculate heart rate with high accuracy[5-7]. There has been a huge growth over the last years in multimedia applications for portable devices like mobile phones. A variety of methods for lossy compression for videos has been developed to manage bandwidth and memory usage, introducing specific artifacts impairing the visual quality as perceived by the end user[8].

This can be used for creating a mobile software capable of real time, on device motion magnification. Eulerian video magnification is used to reveal subtle changes. [9] shows how EVM can be used in real time by subscribing to the incoming Kinect images and applying spatial and temporal filtering. Real time approach used can help in robotics to improving the human-robot interactions as robots don't have a reliable way to detect human in a scene. Various techniques exist to magnify a video. In [10], the author reviews and compares various motion magnification techniques to reveal subtle changes in a video such as Linear approximation method, Phase based video processing, Riesz pyramid for fast phase based video processing and Enhanced Eulerian video magnification. Remote detection of pulse can be helpful and provide physiological assessment without electrodes. The approach discussed in [11] overcomes problems such as artifacts, noise and expenses. It is based on automatic face tracking and blind source separation of the colour channels into independent components. The results we accurate even in the presence of motion artifacts. Spatial decomposition is an important step in EVM. The efficiency and time taken by computation are affected by this step. [12] discusses a new image decomposition technique, Riesz Pyramid. It is suitable for EVM but much less complex and overcomplete than other pyramids. The output motion magnified video has comparable quality but the time taken is very less when compared to other techniques. Riesz pyramid is constructed by breaking the input image into non-oriented sub-bands. Riesz pyramid is efficient because of shared computation of different bands and symmetry of filters.

Laplacian pyramid is used in spatial decomposition to remove pixel to pixel correlation by subtracting a low pass filtered copy of an image from image itself. Further data compression is achieved by quantization. Iteration of the process generates a pyramid structure [13]. Some motions happen without changing their position or without any movement. [14] describes a technique a quadrature pair of oriented filters to vary the local phase, giving the sensation of motion while removing the drawbacks such as jerky display of motion due to snapping and clutter. The cameras are getting faster and faster, but there's a need for conventional frame rate video for human perception. [15] presents a technique with controlled sampling behaviour to convert high fps input video into a conventional speed output video in real time.Fourier transform in temporal domain can be applied to visualise and analyse time dependent effects. Frequency filters are needed to process an image in temporal domain.[16] discusses different frequency filters such as Ideal bandpass filter, Butterworth filter, IIR filter and Gaussian filters. Different filters are needed for different applications in EVM. EVM can be performed while there are multiple subjects in a scene. [17] combines the ability to detect face and amplify them individually to observe each subject's heartrate.This advances the use of EVM by bringing multiple subjects in the same scene and processing them individually. While detecting motion it is important to remove unwanted motion.[18] presents a semi-automated technique for selectively de-animating video to remove large scale motions so that other motions are easier to see.This technique uses graph-cut-based optimization to composite the warped video regions with still frames from the input video. The focus of motion analysis has be estimating the flow vector of the pixel.[19] goes beyond pixel level and uses layers to interact with videos.Layer representation van be used to reveal changes and make them visible to human eyes.[20] presents an efficient technique for video stabilization that achieves high quality camera motion.The approach assembles tracked features in the video into a trajectory matrix, factors it into two low-rank matrices, and performs filtering or curve fitting in a low-dimensional linear space.

Chapter 3

Analysis and Technique

Motion Magnification techniques extract information about the motion that are needed to be observed and after extraction that particular motion is magnified so that it is noticeable and easily observed by naked human eye. Some techniques are based on tracking motion and some are not. The most commonly used motion extraction are:

- Optical Flow
- Spatio-Temporal Filtering

3.1 Optical Flow

Optical Flow is the pattern of relative motion between two objects i.e. visual scene and an observer's eye. Optical flow is used to study a myriad variety of motions— between moving observer and static objects, between static observer and moving objects, or when both are moving. The optical flow field is the velocity field that represents the 3-D motion of object points across a 2-D image. This is actually based on tracking motion.
It is a basic approach involving two steps:

- From one pixel to the other following pixel in each frame we compute its translation.
- Re-rendering the video with small motions amplified.

Also considering a basic approach like this usually leads to artifactual transitions between 2 types of pixels: amplified and un-amplified pixels within

a single structure. Most of the steps of motion magnification are calculated by estimating motions, and also then the pixels are clustered whose motions are then magnified as a group. The actual procedure is to estimate a motion at every pixel, which we begin by analyzing and grouping the motions of feature points and also of local intensity configurations that are made to promise candidates for finding reliable motion trajectories.
It involves following steps:

3.1.1 Register input images

When we began getting frames from the input video sequence, we register them in such a way that small, non-avoidable frames because of camera shake do not get amplified. For this, we take the input sequence as a static scene. We then get the detected feature points and then find refine warp which removes these set of tracked feature points, also ignoring outliers. The resulted registered images are then ready for the further steps.

3.1.2 Cluster Feature Point Trajectories

In order that the motion magnification does not break apart any of the coherent objects, we think to group objects that are usually moved with correlated (not necessarily identical) motions. To achieve this, we roughly track feature points which is present throughout the sequence, then cluster their trajectories into K sets of correlated motions. The background cluster is the one special cluster of feature points that has no translation over frames. An important contribution of registering input images is the computation of trajectory correlation in a manner which is different to the overall scale of the motions, which allows very small motions to be grouped with larger motions to whichsoever they are correlated.

3.1.3 Segmentation: layer assignment

From the previous clustered feature point trajectories, now we wish to derive motion trajectories for each pixel of the reference frame. We interpolate a dense motion field that is for each motion cluster, which gives us K possible motion vectors at each and every pixel. We then need to assign each pixel of each and every frame to one of the layers i.e. clusters or motion layers.
It is possible, in principle, that we can perform segmentation using motion alone, but also we should know that reliable segmentation usually requires the use of additional features. We use so many characteristics of an image-

pixel color, position, as well as motion which helps us in estimating the cluster assignment for each and every pixel, defining a Markov random field which we usually solve using graph cuts. Now to impose temporal consistency, we finally then assign each pixel trajectory to its most commonly assigned cluster with overall time frames. This then gives us a layered motion representation, but also generalises layer membership that includes correlated motions, and not only just the similar ones.

3.1.4 Magnify motions of selected cluster

Subsequent to covering past strides, we find different layers and afterward the client indicates the layer on which a specific movement needs to be amplified. In this amplification all moves are opened up from the reference position by a steady component (range from 4 to 40). Be that as it may, likewise more broad movement amplification capacities are conceivable.

3.1.5 Render video

After the movement amplification steps, we make or render the last video succession. The foundation layer is steady for all edges what's more, thusly we make its pixels first. Presently the pixels of the external layer are replicated as they showed up in the enlisted input outlines. In conclusion, the pixels of the forgotten layers are composed in the yield succession. The forces found are those of the reference outline; the relocations are those of the movements that are measured or amplified, as it is proper to the layer.

3.2 Spatio-Temporal Filtering

Spatio-temporal filtering is one of the most promising approaches to estimate motion. This approach based on continuous wavelet transform (CWT) can be used for motion selectivity. Temporal filtering is used to extract invisible motions. In this method a motion is not tracked.

3.3 Eulerian Motion Magnification

This approach for the motion magnification is motivated by the Eulerian perspective. According to it, certain properties of fluids like pressure and velocity change over time. In this case, the variation of pixel values over time is magnified and studied. In Eulerian approach to motion magnification, motion is not explicitly estimated, but rather exaggerated by amplifying temporal colour changes at fixed positions. It relies on the Taylor series expansions common in differential optical flow analysis. It uses linear approximation.

3.3.1 Spatial Filtering

Spatial processing is the first step in Eulerian motion magnification. Video is decomposed in different spatial frequency bands using pyramid structure. These frequency bands may consist of different spatial frequencies depending on signal to noise ratio (SNR). Later on the amplification is reduced for these bands to suppress artifacts by low pass filtering. Frames of the video are later down sampled to increase computational efficiency. These selective pixel bands are used in temporal filtering to detect motion.

3.3.2 Temporal Filters

After spatial filtering , temporal filtering is done. Here, the time series is considered corresponding to the value of a pixel in a frequency band and a bandpass filter is applied to extract the required frequency bands of interest. The temporal processing is identical for all spatial levels, and for all pixels within each level. The choice of an appropriate temporal filter depends on a video type and information needed to be extracted from it. The filter varies with frequencies we are dealing with. In case of colour magnification the pass band filter specifications are different than in case of small motion magnification. This difference in filter selection is due to frequency bands. If we have to observe the colour changes in face due to blood perfusion in the body, then narrow passband filter is required. Hence ideal bandpass filter with sharp cut of frequencies can be used. To observe small movements of objects broad passband filters are recommended.
Some Temporal Filters that can be used are :

1. **Ideal Bandpass Filter**
   The ideal bandpass filter only allows the frequencies in the passband to pass and gives an output that is mostly ringed or blurred. Ideal bandpass filter is the easiest to stimulated.

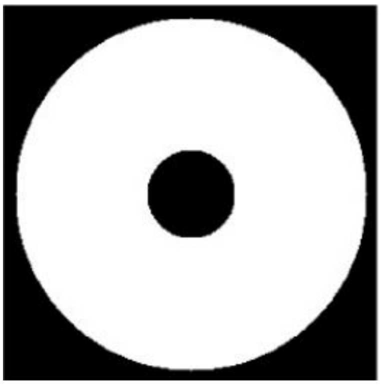

Figure 3.1: Frequency Domain of Ideal Bandpass filter

2. **IIR Filter**
   In real time it is convenient to use IIR filters both for color and motion exaggeration. Two low order IIR filters can be combined to implement a band pass filter.

3. **Gaussain Filter**
   Based on performance Gaussian filter is the best among previously discussed filters because there is no ringing effect in the spatial domain. Gaussian when transformed between spatial and temporal domains remains Gaussian.

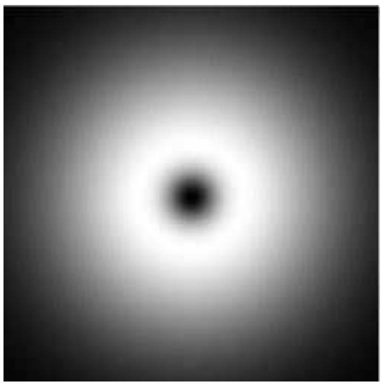

Figure 3.2: Frequency Domain of Gaussian Filter

4. **Butterworth Filter**
   Butterworth filter is designed to have a flat frequency response in the passband. It is also referred to as a maximally flat magnitude filter. They have slower roll off. This filter is mathematically derived

by multiplying the transfer function of high pass and low pass filters. Butterworth filter has a smooth transfer function without clear cut-off frequency or discontinuity.

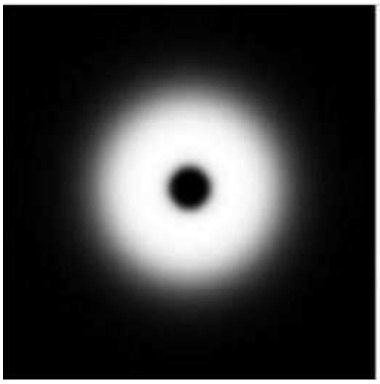

Figure 3.3: Frequency Domain of Butterworth Filter

3.3.3 Amplification factor

After extracting or selecting the required band of frequencies which needs to be amplified like color or magnified as in case of motion, these bands are exaggerated by a factor alpha. These values can be controlled by the user depending upon application requirements. Best value or an optimized value can be obtained by doing experimentation and observing results.

3.3.4 Video Reconstruction

After different frequency bands which have been pooled and temporally processed, these frequency bands are restructured to form a video in which motion is magnified. Video reconstruction is done by collapsing a pyramid used to a spatially decompose image.

Chapter 4

Implementation of Proposed Work

In this Chapter, we will be discussing about the implementation of the proposed work.

4.1 Software Used

To achieve motion magnification, we need to accurately measure visual motions, and group the pixels to be modified. For this we use MATLAB.

MATLAB is a high-performance language for technical computing. It integrates computation, visualization, and programming in an easy-to-use environment where problems and solutions are expressed in familiar mathematical notation. Typical uses include:

- Mathematical computations
- Development of Algorithms
- Modelling, simulation, and prototyping
- Analysis and Visualization of Data
- Graphics Engineering

- Application development, including GUI building

MATLAB is an interactive system whose basic data element is an array that does not require dimensioning. This allows you to solve many technical computing problems, especially those with matrix and vector formulations, in a fraction of the time it would take to write a program in a scalar noninteractive language such as C or Fortran.
MATLAB features a family of "Toolboxes" which serve as application specific solutions for a variety of purposes. Its very important to learn how to use these toolboxes as they allow us to learn and apply specialized Technology. Toolboxes can be understood as huge collections of predefined MATLAB functions that enable the MATLAB enviornment to solve particular types of problems. Areas in which toolboxes are available include signal processing, control systems, neural networks, fuzzy logic, simulation.

4.2 Recording Video

Even though video recording is a simple procedure, we need to keep several things in mind when the video has to be magnified.

- The motion which is not relevant must be minimized. Place the camera on the tripod. If possible, provide and kind of support for the subject (e.g. hand on a table, stable chair).
- Try to make image noise as minimum as possible. It is possible only with the use of a camera with a good sensor, make sure there is enough light so that the ISO can be higher. This ensures the absence of grainy images.
- The subject should occupy the most of the frame and if possible recording should be done in the highest resolution. Better signals can be extracted if only the object of interest is covered properly.
- Uncompressed video should be preferred. Codecs that compress frames independently (e.g. Motion JPEG) are usually preferable over codecs exploiting inter-frame redundancy (e.g. H.264) that, under some settings, can introduce compression-related temporal signals to the video.

4.3 Decomposition of video in Spatial Frequency Bands:

Spatial frequency is actually a characteristic that is periodic across position in space. It tells that how often a structure is repeated per unit of distance. High spatial frequency means that there are abrupt changes in the image such as in its edges and also tells us about it's fine details. Low spatial frequency tells about general orientation. In general, spatial processing comprises of Gaussian blurring and Laplacian Pyramid after which in each of these bands temporal processing is performed. Let us first start with the understanding of the pyramids. We know, Digital image processing is a point of interest in many fields. These days a number of methods are used for enhancement of an image quality and removing unwanted pixels or effects. Image decomposition plays a vital role in making images and videos more prominent and extracts useful information to remove noise from corrupted images. Image pyramids offer a flexible, convenient multiresolution format which makes it convenient to process images.

4.3.1 Image Pyramids

To extract required information from an image is a complicated process. For extracting required information from the image, it needs to be passed through a combination of different filters, with different sequences depending upon the requirements. The visual sense of a human being is very intuitive. It has the properties of less spatio-temporal sensitivity. A detection of variations on color and motions at lower spatial amplitudes is difficult and impossible for the human sense of vision. An image pyramid is used to reveal minor and subtle changes in videos. It is used in gathering important data from an image. An application of image pyramids is very helpful to assemble important aspects and variations in the world around us. Complex mathematical and numerical representation, amplification and decomposition are used to form the image pyramid. A pyramid is a technique developed by computer vision, image and signal processing fields. It is used for signal representation at multiple scales.
There are two types of pyramids:

- **Low Pass Pyramid**
  Here, the first step is to make an image smooth. This process is done with the use of the smoothing filter. The next step is the low pass pyramid and its sub-sampling of the resultant smoothed image from

the results of the previous step. This is mostly done with the factor of double magnitude along each coordinates. This process is repeated and a smaller image is obtained. These results in the images are more smoothened as compared to the previous one. In this scenario, sampling density keeps on decreasing after proceeding 20 of each step in each level. Mathematical and graphical representation of all the processes resembles to a pyramid, thats why it is named as pyramid.

- **Band Pass Pyramid**
  The other type of pyramid is called a band pass pyramid. If we differentiate and form different levels which are adjacent to each other in a pyramid then it is considered as the band pass pyramid representation of image. Although in this process there is an interpolation involved when we represent the adjacent levels of resolution in images.

Now, we come to the most important pyramids used for Image Decomposition.

1. **Gaussian Blur Pyramid**: The Gaussian Blur filter is also known as smoothing filter as removes detail and noise. It uses kernel which is a matrix of numbers. Under convolution, different sizes of kernels will give different patterns of numbers which in turn will produce different results. We know that image is a collection of discrete pixels, therefore we need a proper kernel for the convolution. Once the kernel is selected, the the gaussian smoothing can be done with simple convolution methods. Convolution is done by sliding the kernel over the image, to move kernel all over the image and also not moving it on the sides of the image. The Gaussian outputs a 'weighted average' of each pixel's neighborhood, with the average weighted more towards the value of the central pixel. This helps in removing fine details and also the size of the output image is same as that of the input image. Now this collection of blurred images will help us in making in Laplacian Pyramid, further on which the temporal processing would be performed.

2. **Laplacian Pyramid**: Laplacian pyramid was developed by Burt and Anderson in 1983. The purpose of the research was compression of an image. To attain the high com- pression, image correlation needs to be removed which is done by combining transform coding techniques and prediction. If good compression results are required then the decorrelation of the image has an important impact.

(a) **Structure**

A complete image representation is an important property of the Laplacian pyramid. To recover an original image the process of constructing a pyramid is reversed. The top level of the image L_N , is first expanded then added with L_{N-1}, results in G_{N-1} ,this process continues, and add L_{N-2}, and form G_{N-2}, and so on. It can be written as:

$$G_0 = \sum L_{l,1}$$

Here the Pyramid is used for supporting the scaled image analysis. The levels of Pyramids are formed by some steps which include 'expand' and 'reduce' operations. It is possible with the help of Fast Fourier Transform(FFT). A 20 to 30 bit arithmetic with direct convolution and equivalent weighting function is required to keep a similar accuracy as cascade of convolution with a small generating kernel using 8 bit arithmetic.

The following illustration denotes the comprised steps of Gaussian and Laplacian Pyramid. Here l is the gaussian level and h id the laplacian level.

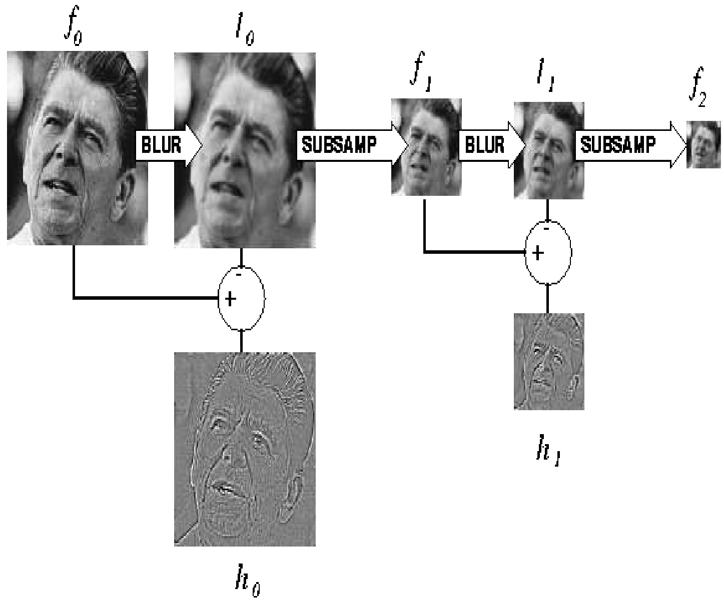

Figure 4.1: Levels of Gaussian and Laplacian Pyramid

The pyramid decompositions for image processing can be defined through the following algorithm. Let the original image be denoted by i and the following blurred and fine images can be denoted by b1, b2,... and f1, f2,... respectively.

- Apply Gaussian filter to i to get our first blurred image b1.
- Now subtracting the blurred image b1 from the original image i. This gives us our first fine image f1, which is showing fine details of the image. Also, if we add f1 and b1, we will get our original image i.
- Now we subsample the blurred image b1 to get blurred image b2 with some aliasing. This b2 is also half the size of the original image i..
- Now to get f2, we interpolate (add zeroes) b2 to full size to subtract it from the original image.

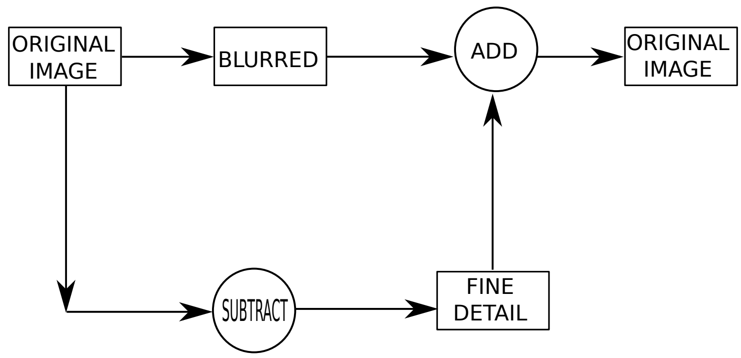

Figure 4.2: Block Representation of Laplacian and Gaussain Pyramids

Therefore, this technique when performed repeatedly by recursively splitting off finer details from the blurred image results in a data structure known as "Laplacian Pyramid".

4.4 Temporal Processing:

Now to perform filtering on each of the spatial frequency bands we need to choose the correct filter according to our application. Here in this project we have worked on the colour change of the face as heart pumps the blood. The heart beat is varied from 40-100 beats per minute and in a normal human being the pulse rate is usually 72 beats per minute. These beats per minute of 40-100 will correspond frequencies within 0.6-1.6 Hz and for normal human being the frequency can be 1.2 Hz. Now we can either use a narrow passband or a broad passband filter. For Motion Magnification we usually prefer broad passband to amplify motions and for color amplification we prefer narrow passband filters as these produce more noise free result. After applying the required filter, we multiply the extracted bandpassed signal by a magnification factor α. This factor is adjustable by the user. The magnified signal is then added to the original laplacian pyramid signal. Then the final step is to collapse the pyramid the generate the final output i.e. the amplified output video.

4.5 Steps in Brief

Our technique is based on Eulerian perspective; in this we can amplify motion without tracking. We use spatial and temporal processing for colour and motion magnification. First spatial filtering is done using laplacian pyramid to decompose video into different frequency bands. Then temporal filtering is performed using same filter on all the decomposed bands to select the passband of interest. The filtered bands are then amplified by a factor alpha to magnify the change. The amplified signal is then added back to the original video to obtain the output. Filters and alpha factor can be changed according to requirement.

4.5.1 Input Video

Input video used is of MPEG-4 format. Input video used is as still as possible to minimise unnecessary motions. The video of face is used to observe the colour changes in face due to blood perfusion.

4.5.2 Spatial Decomposition

The video is first decomposed into pools of different signal-to-noise ratio and frequencies using laplacian pyramid. The goal of spatial decomposition is to increase the signal-to-noise ratio by pooling multiple pixels. Frames of video are pass through low pass filter and downsampled for computational efficiency.

Figure 4.3: Original Image

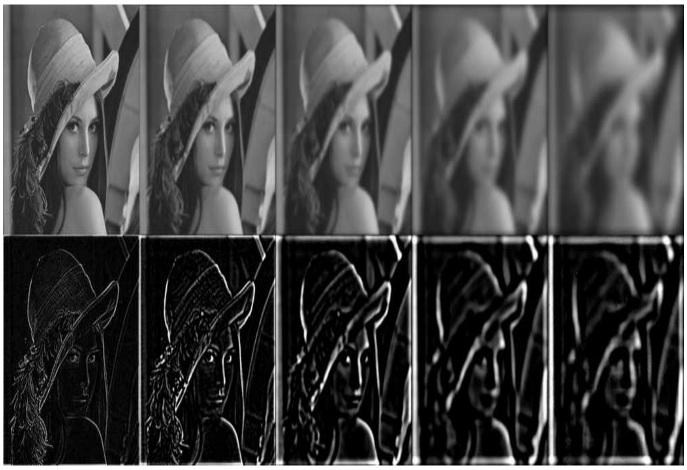

Figure 4.4: Gaussian Stack (above) and Laplacian stack (below)

4.5.3 Temporal Filtering

Temporal filtering is done to extract the frequency band of interest which is needed to be amplified. The sequence of frames is passed through an ideal band pass filter with passband 0.83 Hz (50 bpm) to 1 Hz (60 bpm) to extract the colour change due to blood perfusion. Different filters and passbands can be used according to requirement. The temporal filtering is same for all the levels of the pyramid.

4.5.4 Choice of Filters

One needs to select filter intelligently to extract the desired information. To magnify motion passband has to be large but in case of colour magnification narrow passband is recommended. For colour amplification, ideal bandpass filters are used as they have sharp cut-off frequencies. IIR filters are used for both motion and colour magnification.

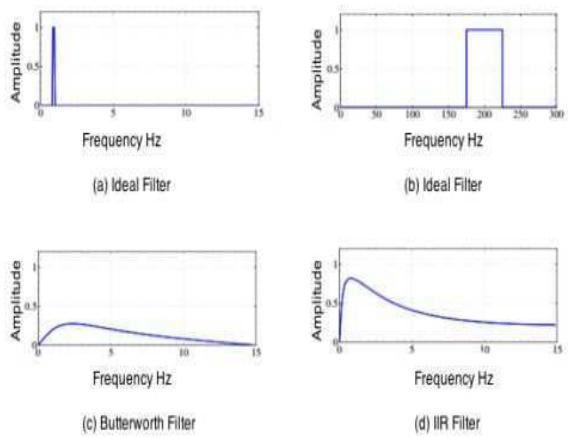

Figure 4.5: Choice of Filters

4.5.5 Amplification Factor

Amplification factor is user defined and is based on requirement. There are some bounds though on amplification factor. A large amplification factor could increase noise and artifacts. A small amplification factor on the other hand may hide necessary information. An alpha=100 is applied to the resulting spatially low pass signal to exaggerate the colour changes due to blood perfusion.

4.5.6 Reconstruction

Reconstruction is done using the reverse laplacian pyramid and pooling the magnified frames together to give the output. The magnified frames are added back to the original frames. The colour change due to blood perfusion was observed.

These steps can be summarized by the following Flow Chart.

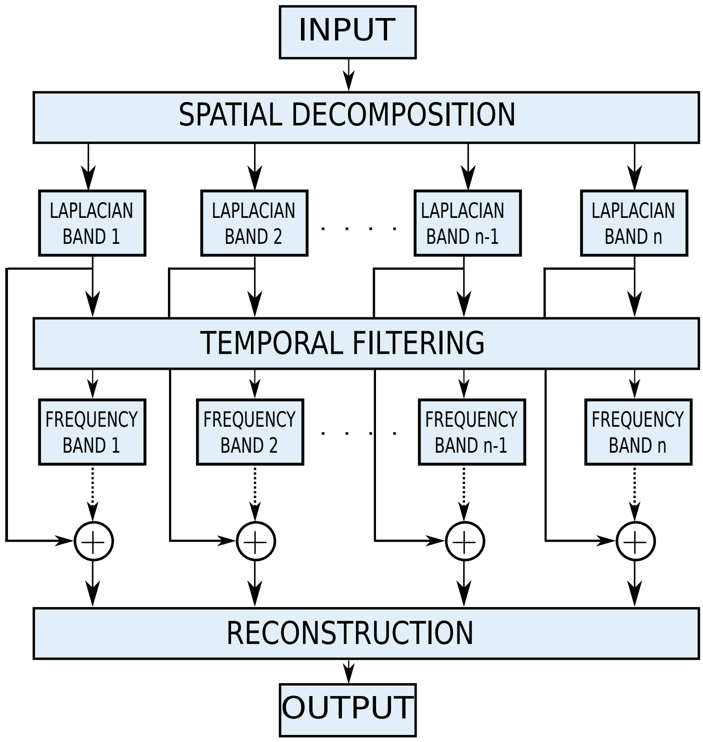

Figure 4.6: Flow Chart of proposed work

4.6 Applications

Motion Magnification has a huge potential and its applications are limitless. We discuss some of the most promising applications of this project.

4.6.1 Medical

The most promising application is real time heart rate monitoring. By observing the rate of flow of blood to the face, also known as blood perfusion, we can measure the heart rate of the subject. This can be very useful in detecting imminent heart attacks and also for health monitoring. Another method of measuring the heart rate is by checking the pulse of the subject where we magnify a video of the subject's wrist and the pulse can be clearly detected without the need of making contact with the subject. In nurseries and hospitals, it is difficult to monitor new born babies. Video surveillance is simply not enough to ensure that the babies are healthy. Motion Magnification can be applied here. A seemingly still video of a sleeping baby will give us a visual representation of the breathing of the child, thus providing highly detailed information. It can also show us the flow of toxic substances in the body. A snake bite injects poison in the body, this venom is not visible to the human eye. but by applying motion magnification, this venom can be traced easily.

4.6.2 Defence and Surveillance

With terrorist activity across the country and daily attacks by intruders at our borders, military and defense is a tough task. We believe that motion magnification can be applied in this field to make management of border security easy. It can act as a visual radar for detection of enemy movements in the visible area. Any tiny movement will be magnified. This works in case of camouflaged enemies too. This visual radar can be modified to detect movements behind smoke, and also at night so as to provide ever more vigilant security. This technique can also be used in shops et cetra for surveillance and security from shoplifters.

4.6.3 Fault Detection

Often, machinery, expensive or cheap can have some inherent defects which may pass testing and checking and may get handed to general public. This defective machinery may cause damage to users and may also cause loss of money and energy. Motion magnification of a running device such as a motor or engine can highlight several flaws in the machinery, hence helping us to avoid those problems. This is particularly useful for testing purposes.

4.6.4 Research and Development

Motion Magnification, being a very new topic hasn't been researched much and very less work on it has been done till date. This makes it a hot topic for R&D and many researchers are working on it and developing new applications based on it.

Chapter 5

Results and Conclusion

5.1 Results

We have taken an input still video and then for different values of amplification factor, we got the output video sequence. The results were generated on MATLAB and also it took few minutes to for the computation of the video. In the output video the color change can be easily seen which is due to the blood flow in our face.
We have taken screenshots of the video for two instants: when the blood flows in the face and when it does not i.e. the face seems to be pale.

5.1.1 Amplification Factor = 50

The following figures shows two frames. The figure (a) depicts the blood flow out of the face back to the heart and the figure (b) depicts the blood flow into the face.

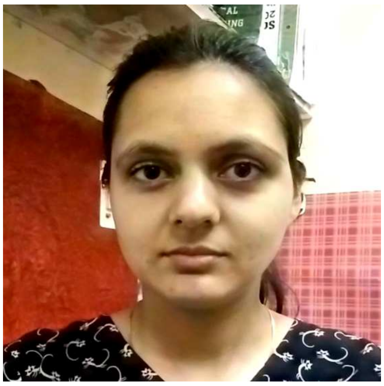

Figure 5.1: (a) Alpha = 50

Figure 5.2: (b) Alpha = 50

5.1.2 Amplification Factor = 55

The following figures shows two frames for the amplification factor 55 and also the colour change is more visible than the previous one.

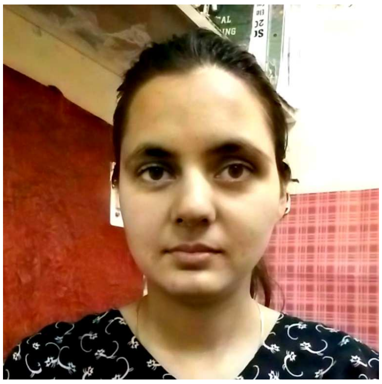

Figure 5.3: (a) Alpha = 55

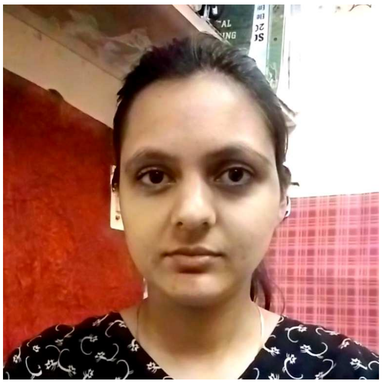

Figure 5.4: (b) Alpha = 55

5.1.3 Amplification Factor = 60

Here we increase the amplification factor to 60. The colour changes are clearly visible but some noise is also amplified.

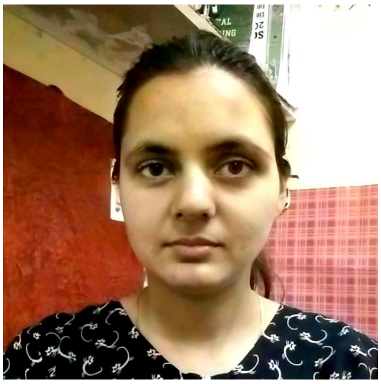

Figure 5.5: (a) Alpha = 60

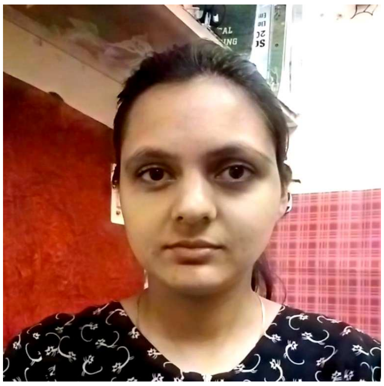

Figure 5.6: (b) Alpha = 60

5.1.4 Amplification Factor = 65

By increasing the amplification factor to 65 the colour changes are exaggerated.

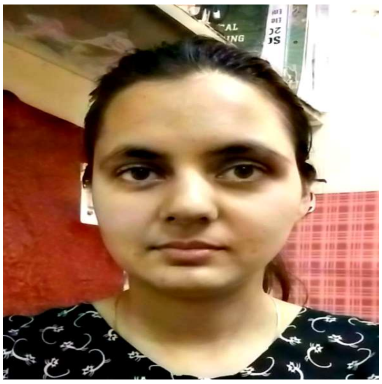

Figure 5.7: (a) Alpha = 65

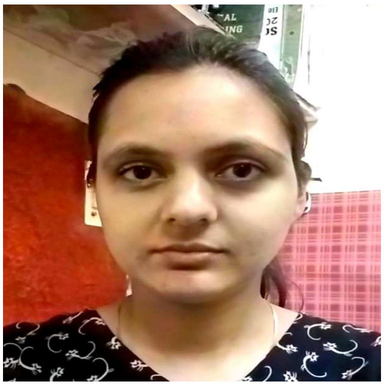

Figure 5.8: (b) Alpha = 65

5.2 Conclusion and Future Scope

We took a normal video as input and magnified it to amplify the tiny motions in it that were invisible to our naked eyes. Gaussain pyramids were implemented in MATLAB to observe these small changes and to produce an output in which these were clearly visible. Motion Magnification works better in some specific cases and applications. In future, experimentation with different filters can help produce better results. Also, this technique has highly accurate results, as can be seen in our results.
It has numerous applications, some of which are discussed by us, such as in security and medical fields.

References

1. Mohamed A. Elgharib, Mohamed Hefeeda, Frédo Durand, William T. Freeman; **Video Magnification in Presence of Large Motions**; Presented in IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2015).

2. Frédo Durand, William T. Freeman, Michael Rubinstein; **A World of Movement**; Scientific American, Volume 312, Number 1, January 2015.

3. Neal Wadhwa, Michael Rubinstein, Frédo Durand, William T. Freeman; **Phase-based Video Motion Processing**; ACM Transactions on Graphics, Volume 32, Number 4 (Proc. SIGGRAPH), 2013.

4. Hao-Yu Wu, Michael Rubinstein, Eugene Shih, John Guttag, Frédo Durand, William T. Freeman; **Eulerian Video Magnification for Revealing Subtle Changes in the World**; ACM Transactions on Graphics, Volume 31, Number 4 (Proc. SIGGRAPH), 2012.

5. Ce Liu, Antonio Torralba, William T. Freeman, Frédo Durand, Edward H. Adelson; **Video Magnification**; ACM Transactions on Graphics, Volume 24, Number 3 (Proc. SIGGRAPH), 2005.

6. Guha Balakrishnan, Frédo Durand, John Guttag; **Detecting Pulse from Head Motions in Video**; 2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

7. Michael Rubinstein, Neal Wadhwa, Frédo Durand, William T. Freeman; **Revealing Invisible Changes In The World**; Science Vol. 339 No. 6119, Feb 1 2013.

8. Shahid, M.; Rossholm, A.; Lovstrom, B.; **A reduced complexity no-reference artificial neural network based video quality predictor**; 4th International Congress on Image and Signal Processing (CISP), 2011, vol.1, no., pp.517,521, 15-17 Oct. 2011.

9. Kiana Alcala, Kathryn Baldauf; **Implementing Eulerian Video Magnification to detect subtle changes in motion**; Jivko Sinapov, CS378,Dec 15,2016.

10. Kranti Kamble, Nitin Jagtap, R.A Patil, Ankit Bhurane; **A Review: Eulerian Video Motion Magnification**; International Journal of Innovative Research in Computer and Communication Engineering,Vol. 3, Issue 3, March 2015.

11. Poh. M. Z., M. Ccduff, D. J., Picard, R. W; **Non-contact, automated cardiac pulse measurements using video imaging and blind source separation**; Opt. Express 18, 10, 10762–10774, 2010.

12. N. Wadhwa, M. Rubinstein, F. Durand, and W. T. Freeman; **Riesz pyramid for fast phase-based video magnification**; 2014 IEEE International Conference on. IEEE, 2014.

13. Burt, P., Adelson, E.; **The Laplacian pyramid as a compact image code**; IEEE Trans. Comm. 31, 4, 532–540,1984.

14. Freeman, W. T., Adelson, E. H.,Heeger; **Motion without movement**; ACM Comp. Graph. 25, 27–30,1991.

15. Fuchs, M., Chen, T., Wang, O., Raskar, R., Seidel, H.-P.,Lensch; **Realtime temporal shaping of highspeed video streams**; Computers and Graphics 34, 5, 575–584, 2010.

16. Leonardo O. Iheme; **Frequency Domain Bandpass Filtering for Image Processing**; EE 583.

17. Matt Estrada,Amanda Stowers; **Amplification of Heart Rate in Multi-Subject Videos**.

18. J. Bai, A. Agarwala, M. Agrawala, and R. Ramamoorthi; **Selectively de-animating video**; ACM Trans. Graph.,31(4):66:1–66:10, 2012.

19. C. Liu; **Beyond Pixels: Exploring New Representations and Applications for Motion Analysis**; PhD thesis, Massachusetts Institute of Technology, 2009.

20. F. Liu, M. Gleicher, J. Wang, H. Jin, and A. Agarwala; **Subspace video stabilization**; ACM Trans. Grap., 30(1):4:1–4:10, 2011.

21. J. Shi and C. Tomasi; **Good features to track**; CVPR, pages 593–600, 1994.

22. Z. Wang, A. Bovik, H. Sheikh, and E. Simoncell; **Image quality assessment: from error visibility to structural similarity**; IEEE Transactions on Image Processing, 13(4):600–612, 2004.

23. A. Joshi, A. Kulkarni, S. Chandran, V. K. Jayaraman, B. D. Kulkarni; **Nadi Tarangini: a pulse based diagnostic system**; Proc. of the IEEE International Conference on Engineering in Medicine and Biology Society Lyon, pp. 2207-2210, August 2007.

24. Y. Chen, L. Zhang, D. Zhang, D. Zhang; **Wrist pulse signal diagnosis using modified Gaussian models and fuzzy C-means classification**; Med. Eng. Phys., vol. 31, no. 10, pp. 1283-1289, December 2009.

25. K. Y. Shin, T. K. Kim, J. S. Song, O. S. Jin; **Remote blood pressure monitoring using a pulse diagnostic system in TCM and**

Android-based Tablet PC; Proc. of the IEEE International Conference on Biomedical and Health Informatics (BHI), pp. 647-650, January 2012.

26. B. Dong, P. H. Gao, H. W. Wang, S. Z. Liao; **Clustering human wrist pulse signals via multiple criteria decision making**; Proc. of the IEEE International Conference on Tools with Artificial Intelligence, pp. 243-250, November 2014.

27. D. Y. Zhang, D. Zhang, D. Zhang, Y. P. Zheng; **Wavelet based analysis of Doppler ultrasonic wrist-pulse signals**; Proc. of International Conference on BioMedical Engineering and Informatics, pp. 539-543, May 2008.

28. A. Alzahrani, A. Whitehead; **Preprocessing realistic video for contactless heart rate monitoring using video magnification**; Proc. of 12th Conference on Computer and Robot Vision (CRV), pp. 261-268, June 2015.

29. P. Burt, E. Adelson; **The laplacian pyramid as a compact image code**; IEEE Transactions on Communications, vol. 31, no. 4, pp. 532-540, April 1983.

30. X. He, R. Goubran, X. Liu; **Using Eulerian video magnification framework to measure pulse transit time**; Proc. of IEEE International Symposium on Medical Measurements and Applications (MeMeA), pp. 1-4, June 2014.