# Automation of Region Specific Scanning for Real Time Medical Systems

Denis Wong
University of Cape Town
Cape Town, South African
Email: denis.wong@uct.ac.za

Fred Nicolls
University of Cape Town
Cape Town, South Africa
Email: fred.nicolls@uct.ac.za

*Abstract*— **X-rays have played a vital role in both the medical and security sectors. However, there is a limit to the amount of radiation a body can receive before it becomes a health risk. Modern low dose x-ray devices operate using a c-arm which moves across the entire human body. This paper shows how radiation can be reduced on a human body by isolating the region that requires exposure. This work is based on a medical scanner that is still under development and therefore a prototype of the scanner is developed for running simulations. A camera is attached onto the prototype and used to point out the regions that are required to be scanned. This is both faster and more accurate than the traditional method of manually specifying the areas, as it also accommodates minor movements from the patient. An analysis is performed on the automation process as there are many variables such as speed, accuracy and searching thresholds that need to be catered for in the experiment. It is found that the correct region of interest can be located with the use of reliable feature points and that certain regions of the body are easier to locate than others. Currently, partial scans are done manually and this is a step forward towards automating the process completely.**

## I. INTRODUCTION

Digital image processing has become a field of growing interest, especially in industries such as the medical sector. Digital image processing is achieved by a set of computer algorithms to perform image processing on digital images [1] to aid in the analysis of the human body. Image processing can be in the form of analysis or manipulation on digital images.

Lodox Systems designed and developed its medical x-ray scanners which originated from the Scannex [2], an x-ray security scanner developed by De Beers, to prevent diamond theft within the mining industry [3]. The Lodox scanners are unique in that they can produce a full body x-ray image within minutes after a single 13 second scan. Lodox Systems currently has a medical scanner in the market, called the Statscan, and is in the process of developing their latest medical device, the Versascan. The Versascan is designed to be a multi-purpose, self-contained and transportable digital radiography system for general and orthopedic radiography. It is a vertically-orientated scanner, so the patient is also vertically positioned. The apparatus used to capture data needs to mimic that of the Versascan as closely as possible, for experimental purposes. A simple garage motor track, vertically-orientated, is used to capture video data of subjects, with the aid of an attached camera. For the remainder of the paper, the garage motor track is referred to as the c-arm unless otherwise specified.

The Versascan can perform partial scans of the human body, but these have to be done manually. Performing partial scans manually leads to human errors when having to specify the starting and stopping points, using the laser, which is built in the c-arm, as a guide and requires that the patient stand very still during the scan. This paper aims to provide the information needed in order to perform a partial scan of the human body. This proposed setup uses the camera to capture a full body image of the patient during a pre-scan where x-rays are disabled. The operator uses the full body image to mark the region that requires scanning. Using a camera to locate the region of interest, in real-time, provides an allowance for the patient to move slightly. Therefore the main aim of this paper is to find whether it is possible to perform a partial scan, on the Versascan, with the aid of a camera. It is proposed that the camera be attached to the c-arm, above the x-ray source.

There are two important practical aspects, namely the workflow for the radiographer and the processes that occur to locate the marked region. Figure 1 is a flowchart containing both aspects for performing a partial scan automatically. The elliptical elements show the steps taken by the radiographer and the rectangular elements indicate what processes are performed, at each step, in order to locate the region of interest automatically. The workflow for the radiographer is important if the proposed modification is to be accepted for the Versascan design. The reason is because the radiographer needs to know what the procedures are in order to perform a partial scan and locate the region of interest automatically.

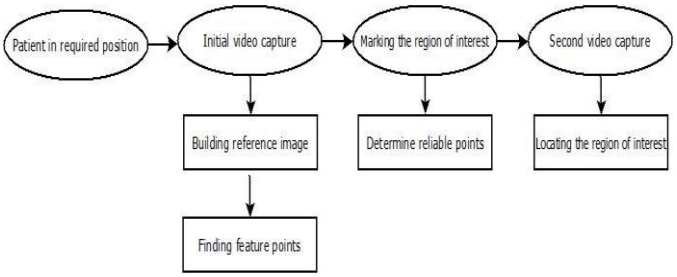


Figure 1: Flowchart of performing a partial scan automatically.

This approach requires that there are pairs of video data in order to perform the relevant tests, one for the reference image and one for the scanned image to search for the region of interest. The pairs of video data are of patients in standard poses with slight adjustments to their stances so as to mimic minor movements in a realistic situation.

Section II looks at relevant academic research from various papers, articles, websites and books in the image processing field. A discussion of the entire proposed workflow for the radiographer and the processes to perform a partial scan automatically is given in section III. All assumptions and experiments performed are evaluated in section IV. Section V presents a summary of the entire paper with its findings, followed by a discussion of ways in which this work could be taken further.

## II. RELATED WORK

### A. Template Matching

Template matching is commonly used for object recognition and stereo-matching. Template matching techniques compare portions of images against one another. Correlation values are calculated in the various positions that indicate how well the template matches the image. Correlation is a measure of the degree to which two variables agree, not necessarily in actual value but in general behaviour.

A common way to calculate the position $(u, v)$ of the object in the search window is to evaluate the cross correlation coefficient value, *c*, at each point $(u, v)$ for function *f* and the template *t*. The positions $(u, v)$ represent the shift of *u* in the x-direction and by *v* in the y-direction. Cross correlation is motivated by the Euclidean distance which is a measure of similarity and is shown as

$$d(u,v) = \sqrt{\sum_{x,y}[f(x,y) - t(x-u, y-v)]^2}. \quad (1)$$

Euclidean distance is only appropriate for data measured on the same scale as no adjustments are made for differences in scale. Equation 2 is made by expanding *d* and is shown as,

$$d^2(u,v) = \sum_{x,y}[f^2(x,y) - 2f(x,y)t(x-u, y-v) + t^2(x-u, y-v)]. \quad (2)$$

In equation 2, if $f(x,y)$ and $t(x-u, y-v)$ are standardized, the sums are both equal to a constant value n. Therefore, $\sum f(x,y)t(x-u, y-v)$ is the only non-constant term just as it is in the reduced formula for the correlation coefficient:

$$c(u,v) = \frac{n\sum_{x,y} f(x,y)t(x-u,y-v) - (\sum_{x,y} f(x,y))(\sum_{x,y} t(x-u,y-v))}{\sqrt{(n\sum_{x,y} f^2(x,y) - (f(x,y))^2)(n\sum_{x,y} t^2(x-u,y-v) - (t(x-u,y-v))^2)}}. \quad (3)$$

Lewis states that there are a few disadvantages to using the cross correlation coefficient for a measure of similarity [4]. Some of the disadvantages mentioned are that the range of the correlation coefficient value is dependent on the size of the feature and that it is not invariant to changes in scale and lighting conditions. Lewis states that the difficulties with the cross correlation can be overcome by normalizing the image to unit length [4]. The normalized cross correlation is shown as

$$\gamma(u,v) = \frac{\sum_{x,y}[f(x,y) - \bar{f}_{u,v}][t(x-u,y-v) - \bar{t}]}{\sqrt{(\sum_{x,y}(f(x,y) - \bar{f}_{u,v})^2)(\sum_{x,y} t(x-u,y-v) - \bar{t})^2)}} \quad (4)$$

where $\bar{t}$ is the mean of the feature and $\bar{f}_{u,v}$ is the mean of $f\ (x,y)$. Normalized cross correlation is a popular measure of similarity as its easy hardware implementation makes it useful for real-time applications. Work has been done on increasing the performance of normalized cross correlation with the use of basis functions. Briechle and Hanebeck proposed using rectangular basis functions where the number of calculations depend linearly on the number of basis functions used [5]. The specific example used in [5] has an outcome that results in a computational reduction of 47 times using basis functions.

There have been some image matching methods performed based on normalized cross correlation [6, 7, 8]. However, these methods do not perform well as normalized cross correlation is not invariant to rotation. Zhao, et al. propose a hybrid method, consisting of both feature points and templates, to improve the results of normalized cross correlation [9]. The hybrid method consists of using feature points on the two images to determine the rotation and scale changes according to the characteristic scale and dominant direction of the points. The invariant normalized cross correlation is then applied at the corresponding feature points.

The main difference between the works of [9] and [5] is that the one potentially eliminates the measure's variance and the other increases its performance by reducing its computation time.

Another application for template matching is not only to classify an object but also to track it. Object tracking is usually categorized into two classes. One is where tracking takes place while the camera is stationary and the other is when the camera is moving. The most important characteristic is that to make a real-time system, the image captured by the camera must be processed before the next frame is digitized.

Pal and Biswas propose an automated correlation based tracking approach using edge strength and Hausdroff Distance Transform (HDT) technique for tracking moving targets [10]. The approach produces a complete real-time video tracking system for both detecting and tracking moving targets from optical image sequences.

Other methods used for detecting objects can be seen in [11] and [12] with differences being that the object's shape is known. Cole et al. describe how a 2D model can be used [11]

and Gupta et al. show how detecting objects can be done with a 3D model [12].

### B. Feature Matching

The feature-based matching approach is the easiest method for finding image displacements. The method finds features in an image, such as edges and corners, and calculates the change in distances of the position of the feature points from the original image to another.

When video data is considered, the displacement is calculated from frame to frame. This is basically a two-step approach. Firstly, feature extraction is performed on two or more consecutive frames, to both reduce the amount of information to be processed and to obtain a higher level of understanding of the image scene. Secondly, these feature points are matched between frames to find any change in the positions of the points. Generally, changes in feature point positions between frames usually mean a movement of some object or background.

Feature-based matching is usually preferable when an image has strong features, such as sharp corners, in it. A feature-based approach is generally faster than a template-based approach because it does not consider the entire image but rather only the feature points found.

Edges indicate boundaries in an image, which makes them important for image processing. Edges in an image usually appear as intensity changes in pixels situated next to each other. There are many different methods of edge detection but they can be grouped into two categories, gradient-based and Laplacian-based. Mlsna and Rodriguez show that the difference between the two categories is that the gradient methods consider maximum values in the first derivative of an image and Laplacian methods look for zero crossings in the second derivative of an image [13].

It is shown in [14] that, in practice, a zero crossing filter is created by performing Gaussian smoothing followed by Laplacian filtering. The Laplacian-of-Gaussian (LoG), which is the convolution mask of the zero crossing operator, can be obtained from using various orders of linear filters and the rotational symmetry of Gaussian filter.

Few examples of other edge detectors are Sobel, the Canny, the Local Threshold and Boolean Function Based edge detectors [15] and color edge detection using euclidean distance and vector angle [16]. Nadernejad et al. have performed a greater analysis of various edge detectors, including the ones previously mentioned [17].

Corners are the intersections of two edges of sufficiently different orientations. Therefore corners contain two dimensional features and can potentially represent object shapes. The ability to represent object shapes play important roles in matching and pattern recognition.

There are many different corner detectors that exist such as the Principal Curvature-Based Region (PCBR) detector [18] and the Harris operator [19]. Corner detectors have many applications in motion tracking, stereo matching and image database retrieval. Mokhtarian and Suomela modify the corner detector to make it more robust, based on the curvature scale-space (CSS) representation [20]. The quality of a corner detector is determined by its ability to detect the same corner in multiple images of the same scene but under different conditions, like lighting, translation, rotation and other transformations.

The Harris corner detector is a good method to use to detect corners as it provides good quality corners under varying rotation and illumination and may detect interest points other than corners. For the purpose of this paper the Harris corner detector is considered due to its strong invariance to rotation, scale, illumination variation and image noise [21].

Harris and Stephens propose combining the corner and edge detector based on the local auto-correlation function to obtain feature points for tracking algorithms [22]. Weijer, et al. propose combining the two detectors by photometric quasi-invariants [23] and Ando by gradient covariance [24]. Parks and Gravel provide a detailed comparison of over 10 various corner detectors including ones mentioned previously [25].

Lowe developed and published the algorithm called Scale Invariant Feature Transform (SIFT) to detect and describe local features in images [26]. The University of British Columbia has patented this algorithm but it is available to the public for research purposes only and there are papers available by Lowe that give a better understanding of the SIFT keypoint detector method [27, 28, 26]. The SIFT algorithm is robust because, as the name suggests, it is able to handle image transformations like scale, rotation and deformation. There are four steps that SIFT goes through to transform image data into scale invariant coordinates relative to local features [26]. Aly has shown that SIFT can be used to find feature points in a face to identify a person for surveillance and access control [29]. There are various other applications that use SIFT such as image stitching [30], video tracking [31] and 3D modeling [32, 33].

The advantages and disadvantages of the two matching methods are mentioned in this section in order to get a better understanding of them. A good understanding of the matching methods is necessary in order to create an accurate and reliable online search to locate a region of the body.

## III. Workflow Processes

To achieve a better understanding of the entire proposed system, the methods used are broken up into separate processes. Figure 1 shows the radiographer's workflow where the respective processes are performed at each stage.

### A. *Reference Image*

The reference image is the first image displayed on the workstation and this is where the radiographer marks the region of interest. The reference image is obtained during the c-arm's first pass by stitching the initial video captured from the camera at 60 *fps*. Figure 2 shows examples of stitched reference images. Routine views are generally in the anteroposterior and lateral positions. Figures 2a and 2b show examples of the anteroposterior position and figures 2c and 2d show the lateral position.

(a) (b) (c) (d)

Figure 2: Examples of reference images. (a) and (b) Anteroposterior position. (c) and (d) Lateral position.

The video captured during the first pass consists of 780 frames. The reference images shown in this paper are constructed from video data consisting of 1080 frames. The reason for the additional 300 frames, or 5s, is due to the delay between operating the camera and the garage opener. It was found that the time taken for the workstation to configure the camera and begin capturing was inconsistent. To overcome this, a delay was implemented, such that after 3s of capturing a signal was sent to start the garage opener. The remaining 2s is used to cater for the approximate time taken for the slide to reach the other end.

Image stitching is the process in which multiple images are aligned by various registration algorithms and blended together in a seamless manner [34]. The video datasets were captured by a camera mounted on a vertically-orientated garage motor track. The reference image was created by taking a number of rows, r, at each frame captured. The rows that are used for stitching the reference image are at the centre of every frame. The method of stitching performed uses a number of rows at each frame and stacks them underneath or above each other, depending on the scan direction, i.e. pass 1 or pass 2. The result of taking into account the centre rows of the video data produces the reference image.

By using the centre rows of every frame, it is effectively providing the information that is directly in front of the camera. However, for the second pass, the camera would have to 'look-ahead' to identify the area before the c-arm reaches it. Using the proposed camera configuration, a maximum look-ahead distance of 179.13mm and 435.13mm is achieved if the c-arm travels downwards and upwards respectively for its second pass. Therefore, a recommendation is made that the c-arm's first pass start from the top moving downwards and from the bottom moving upwards for its second pass.

The method of image stitching mentioned causes some concern for data loss as it simply takes a number of rows at the centre in each frame and constructs the reference image from that. It is seen that by considering two rows at a time, a loss of only 0.16% is obtained which is seen to be minimal.

### B. *Finding Feature Points*

Feature points are important as they are used to locate the region of interest on the scanned image. However, one problem was found with this approach: when there is ambiguity or patterns present around the region of interest, corners are sometimes found at other locations that are visually similar. To remove corners that are either ambiguous or found in patterns, a need for more reliable feature points is necessary.

### C. *Determining Reliable Points*

The process of determining reliable feature points occurs after the radiographer has marked the region of interest, as only the feature points which fall within the marked region are considered and the rest are ignored. Only feature points within the marked region are considered as these are the ones used to locate the region of interest on the scanned image.

The approach for determining whether a feature point is reliable or ambiguous consists of looking at the neighbourhood of each point. A small window of size 15 × 15, centred at the detected corner, is considered. Template matching is then performed over the surrounding area of size 75 × 75 to see whether there are other locations which have similar appearances. Normalized cross correlation is selected to measure how similar the feature point is to the background. Various thresholds have an impact on the resulting correlation value which determines how reliable each corner is and this is discussed in more detail in the next section.

### D. *Locating the Region of Interest*

After the region of interest has been marked, the radiographer controls the c-arm to perform the second pass. The second set of video data is not only being stitched together, but also being used to locate the region of interest.

In order to find the region of interest and provide the location to the x-ray source before it passes it, a search is required to take place ahead of the c-arm. Searching ahead of the c-arm is done by adjusting the stitching method during the second pass. Instead of using two rows at the centre of every frame, higher rows are considered. The look-ahead distance is not set to a constant value but is varied depending on the height of the marked region. In the case where the height of the marked region is greater than the maximum look-ahead distance, the maximum look-ahead distance is then considered.

An online searching method is necessary in order to locate the region of interest in the scanned image efficiently and

accurately. The time taken to locate the region of interest is important as it is necessary to identify the marked area before the c-arm reaches it.

The time available for the online search is catered for with the varying look-ahead distance. In order to achieve accuracy within 2% source to image detector distance (SID), the actual region scanned as a result of the search needs to fall within a distance of 20mm from the region marked by the radiographer.

The approach is to use reliable points found within the region of interest, on the reference image, in order to identify the respective area in the scanned image. A template of size 15 × 15 pixels centred on each reliable point is used for the search in the scanned image. To cater for minor movements, a search window of size 51 × 51 pixels is used to provide an allowance of minor movements of 25 pixels in any direction. One factor that determines the size of the search window is the accuracy as a distance error of more than 20 pixels is greater than 5%, which is regarded as a failed test.

The search for the matches for the reliable points on the scanned image yields normalized correlation coefficient values at each point within the search window. The point with the highest match value is regarded as the best match. If the highest match value is greater than some search threshold, then that point is considered a reliable match. If it is below the threshold then the corresponding match is determined to have not been found and the match pair is thus ignored. Once a specified number of corresponding reliable points are found, an estimate of the marked region on the scanned image can be calculated.

The estimated location of the marked region on the scanned image is calculated with the use of the pixel coordinates of both the original and the corresponding match points found. First, the pixel distances are measured, both horizontally and vertically, between each reliable point and the marked region on the reference image. These distances are then transferred to the corresponding match points and are used to calculate the location of the marked region on the scanned image. In principle the marked region can be found by using the pixel distances measured on any single reliable match as they all indicate the location of the marked area.

Figure 3 shows the reference image, on the left, and the scanned image, on the right, as the c-arm moves downwards and performs an online search. In this case, the number of corresponding reliable points is 2. The yellow line indicates the position of the c-arm and the green line indicates the camera's viewpoint, which is also the last row that has been stitched. A closer look at the scanned image is needed to see where the estimated location of the marked region is. Figure 4 shows the same scanned image from figure 3 with the addition of the red box which indicates the estimated location of the marked region.

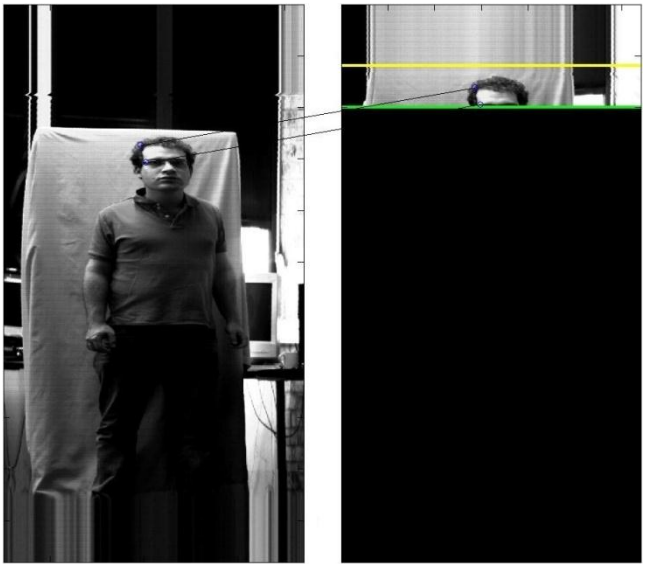

Figure 3: Result of online search after finding two corresponding reliable points. The yellow and green lines indicate the position of the c-arm and the camera's viewpoint respectively.

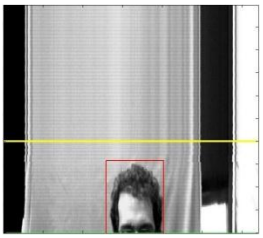

Figure 4: Result of online search indicating the location of the marked region on scanned image. The yellow and green lines indicate the position of the c-arm and the camera's viewpoint respectively. The red box is the estimated location of the marked region.

## IV. Experiments and Results

This section provides a detailed analysis of the experiments performed on the datasets acquired for this paper.

### *A. Ground Truth*

A measure of accuracy needs to be defined for the estimated location of the region of interest on the scanned image. A maximum error of 2% of the SID is allowed in order for the proposed modification of attaching a camera onto the c-arm to be accepted for the Versascan. The ground truth is only used as a measure of accuracy to see how well the search method performs.

Visual inspection can be used to see whether the estimated marked area has captured the required body region, but this doesn't provide a quantified accuracy measure. Therefore, once the second video has been captured, another search is performed. However, in this case instead of doing a progressive search, all of the reliable points on the reference image are used. The region found using all the reliable points, referred to as the ground truth, is then compared against the estimated marked area for an accuracy measure. The ground truth is assumed to be the closest location to the original marked region. In addition to the ground truth, a visual test is

also made to determine whether the correct marked region is found on the scanned image. The distance between the ground truth and the estimated region is the distance error used to determine accuracy in pixels.

Using the previous example where figure 4 shows the estimated location of the marked region on the scanned image, the ground truth is determined and shown in figure 5. Figure 5 shows the entire scanned image where the red and green boxes represent the estimated locations of the marked region and the ground truth respectively.

Using figure 5 as an example, the error is found to be 7 pixels horizontally and 4 pixels vertically which is equivalent to 17.92mm and 10.24mm. Therefore, the example illustrates a successful test as it resulted in locating the region of interest correctly within 2% accuracy. The results of each test in the experiment are analyzed using the ground truth to determine whether the correct region has been found and to what accuracy it is.

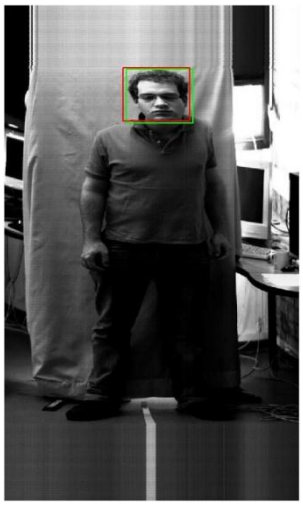

Figure 5: Entire scanned image with an estimated location of the region of interest and ground truth indicated by red and green respectively.

*B. Thresholds*

Various thresholds were mentioned in the workflow processes, all of which influence the results in some way. An evaluation is made on the different thresholds and a range of optimum values is identified that give a suitable result. The three thresholds evaluated are the number of reliable points found on the scanned image, the correlation coefficient value when searching for the reliable points, and the reliability measure of the feature points used.

The experiment consists of two hundred and sixty tests where different regions of the body were marked and searched for. The results in this section have been captured by repeating the experiment and changing the various thresholds accordingly. For experimental purposes, tests which have achieved an error within 2%, 3% and 5% are recorded as passed tests as the marked area identified on the scanned image contains the body region.

Each threshold is varied and a recommendation is made based on two results, the percentage of tests passed and the percentage of those passed tests that are within 2% and 3% error. The outcome of combining the two results is a percentage of tests passed within a certain accuracy. Therefore, a recommendation is made for each threshold based on the combination of the two results. For all the accuracy illustrations, the red and blue points show passed tests within 2% and 3% accuracy respectively.

The number of matches required is an important parameter in the matching process. If this parameter is set to be one, the estimated location of the region of interest would be obtained from the relative position of a match of one point. This makes the process of finding the estimated region fast but potentially inaccurate. On the other hand, if the parameter is set too high then the estimated region might not be found because the number of matches required within the marked region might never be obtained

Therefore, the varying number of matches required used for the experiment are 2, 5, 10, 20 and 50. This particular experiment used the search and point reliability threshold of 0.9 and 10 respectively. It is found that the greater the number of matches required, the more tests that fail. Figure 6a shows the results of the range of the number of required matches considered. The results suggest that one should use a low number of matches for the search. However, figure 6d shows an increase in obtaining more accurate results as the number of required matches increases to approximately 20. The product of combining the two results are, in order: 0.225, 0.273, 0.168, 0.092 and 0.004. Therefore a recommendation is made to set the number of matches required to 5 to cater for both correct and accurate results.

The correlation coefficient values from the search must be greater than the search threshold for a match to be declared. The highest correlation coefficient value around the search window is then used as the best match location to where the matching point is. The search threshold is therefore an indication of how good a match has to be for it to be considered reliable. The thresholds used for this experiment are 0.8, 0.85, 0.9, 0.95 and 0.98. This experiment used a required number of matches and a point reliability threshold of 5 and 10 respectively.

Figure 6b shows the results of the experiment where an increase in the search threshold results in a decrease in the percentage of tests passed. The accuracy of the tests passed is not drastically affected by the varying search threshold, as shown in figure 6e. The product of combining both results are, in order: 0.257, 0.271, 0.273, 0.210 and 0.136. Therefore, a recommendation is made to use a search threshold of 0.9 to cater for both correct and accurate results.

A test is performed on each feature point individually to specify whether it is reliable or ambiguous. This test uses

template matching of size 15 × 15 and a search window of size 75 × 75 with each feature point as its centre, to determine whether there is a similar point nearby. Normalized cross correlation is used as the template matches around the search window.

All the correlation coefficient values are evaluated and accumulated as a weighting to how reliable the point is. If the correlation coefficient is 1, this is generally the case where the template is in its original position and therefore ignored. If the correlation coefficient is greater than 0.95, it is assumed that there is a similar template in the search window and 5 is added to the accumulated weighting. If it is greater than 0.9, then only 1 is added as it is not strongly similar. If the highest correlation coefficient value in the search window is less than 0.9, it is assumed that there are no points similar and ignored.

The accumulated weighting value is then compared to the point reliability threshold. If the weighting is smaller than the point reliability threshold then it is identified as a reliable point. If the weighting is greater than the point reliability threshold it is identified as an ambiguous point and is ignored when performing a search for the marked region. In the experiment, the point reliability threshold values considered are 10, 20 and 50.

To see the effects of the point reliability threshold in the experiment, the required number of matches and search threshold has been set to 5 and 0.9 respectively. It is shown in figure 6c that a higher point reliability threshold results in an increase towards the percentage of tests passed. However, reliable points achieve better accuracy than non-reliable points, as shown in figure 6f. The product of combining the two results are, in order: 0.290, 0.250 and 0.226. Therefore, a recommendation is made to use a point reliability threshold of 10 to cater for both correct and accurate results.

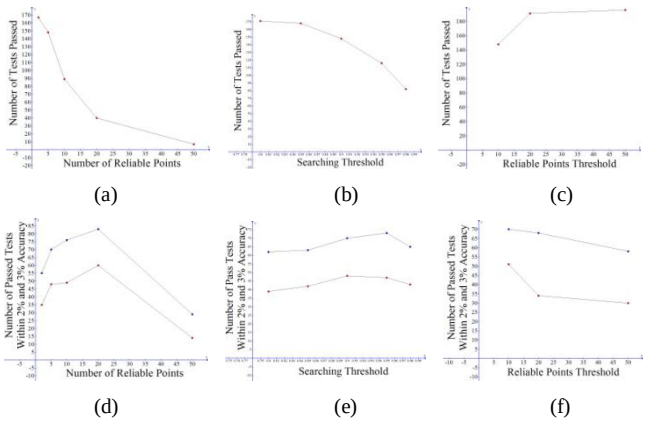


Figure 6: Results of experiment varying various thresholds. (a), (b) and (c) shows tests passed with varying various thresholds. (d), (e) and (f) show the corresponding tests passed within 2% and 3% accuracy.

### C. *Performance of Different Body Regions*

An evaluation is performed on each body region individually to see if some regions are found more easily than others. Recommended values for the different thresholds, mentioned previously, are used in determining the performance of different body regions.

An analysis is performed on each body region to see whether some regions perform better than others. Figure 7 shows the percentage of tests passed for each region.

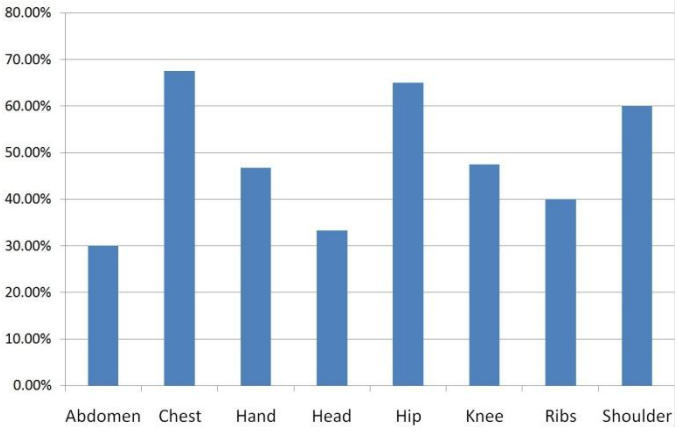


Figure 7: Results of tests passed for each body region.

The poorer performing regions, the abdomen, head and ribs, have been excluded to observe how it affects the overall performance. Figure 8 shows the performance using all the body regions and the other excluding the poorer performing regions. The blue, red and green indicates the tests passed and accuracies within 2% and 3% respectively. Removing the poorer performing regions results in an increase in the number of tests passed without having an impact on accuracy. This shows that certain regions of the body are easier to locate than others using the proposed online search.

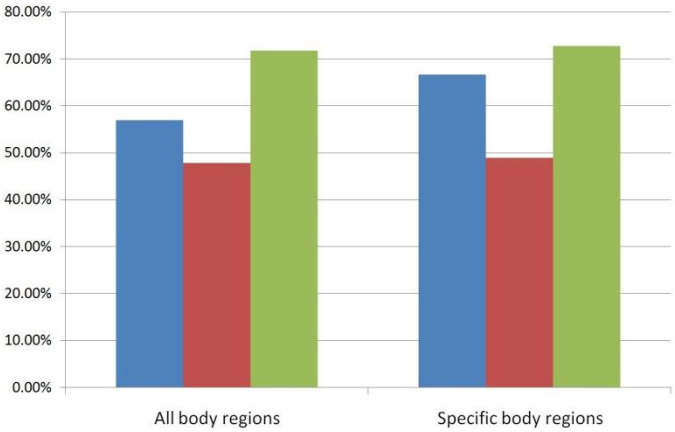


Figure 8: Results of the experiment using all and only specific regions.

## V. CONCLUSIONS AND FUTURE WORK

The most significant result is that it is possible to automate the search for a region of interest on a real-time medical scanner. After performing an experiment consisting of 260 tests, it has been found that it is possible to locate a region on the body, marked by the radiographer, with the aid of a camera attached to a c-arm.

The main factors that influence the search were the thresholds placed on the number of matches required, the search threshold, and the reliability of the feature points. An evaluation on the various thresholds, which consisted of varying the threshold values, was performed in order to see the impact on the results and a recommended value was provided for each threshold. Taking the recommended threshold values into consideration, the results are found to have an overall performance of 57% of which 48% and 72% were within 2% and 3% accuracy.

It was also found that certain regions of the body were easier to locate than others. When ignoring the regions that were harder to locate, the abdomen, chest and head, the overall performance had increased to 67% of which 49% and 73% were within 2% and 3% accuracy respectively. This shows that certain regions of the body are easier to locate than others.

Actual data could not be acquired as the Lodox Versascan is still under development. This was overcome by mimicking the Versascan environment as closely as possible and obtaining datasets accordingly. Once the Versascan is operational, it would be of interest to acquire datasets from the actual device and compare them to the results found in this paper. As this experiment has shown, the method used to perform an online search to locate the region of interest is moderately successful.

## REFERENCES


[1] I. Pitas. *Digital Image Processing Algorithms and Applications.* John Wiley and Sons, Inc., 2000.

[2] De Beers. SCANNEX X-ray Body Scanner, 2011.URL *http://www.debeersgroup.com/ImageVault/Imagesid_1893/scope_0/ImageVaultHandler.aspx*

[3] HERCA Working Group 2. Facts and Figures Concerning the use of Full Body Scanners using X-rays for Security Reason. Oslo *HERCA Plenary Meeting*, June 2010.

[4] J.P. Lewis. Fast Normalized Cross-Correlation. Visual Interface, Canadian Image Processing and Pattern Recognition Society, pages 120-123, 1995

[5] K. Briechle and U.D. Hanebeck. Template Matching using Fast Normalized Cross Correlation. In *Proc. SPIE 4387*, pages 95-102, 2001

[6] W. Forstner. A Feature-Based Correspondence Algorithm for Image Matching. *International Archives of Photogrammetry and Remote Sensing*, 26(3):150-166, 1986.

[7] Z. Zhane, R. Deriche, O. Faugeras and Q.T. Luong. A Robust Technique for Matching Two Uncalibrated Images Through the Recovery of the Unknown Epipolar Geometry. *Artificial Intelligence Journal*, 78(1), 1995.

[8] M. Pilu. A Direct Method for Stereo Correspondence Based on Singular Value Decomposition. *Computer Vision and Pattern Recognition*, pages 261-266, 1997

[9] F. Zhao, Q. Huang and W. Gao. Image Matching by Normalized Cross-Correlation. In Proc. IEEE International Conference on Acoustics, Speech and Signal Processing ICASSP2006, 2:14-19, 2006

[10] S. Pal and P.K. Biswas. Modified Hausdorff Distance Transform Technique for Video Tracking. *Indian Conference on Computer Vision, Graphics and Image Processing*, 2000.

[11] L. Cole, D. Austin and L. Cole. Visual Object Recognition using Template Matching. *Australian Conference on Robotics and Automation*, 2004.

[12] N. Gupta, R. Gupta, A. Singh and M Wytock. Object Recognition using Template Matching, 2008. URL *http://www.stanford.edu/class/cs229/proj2008*.

[13] P.A. Mlsna and J.J. Rodriguez. Gradient and Laplacian Edge Detection. *The Essential Guide to Image Processing (Second Edition)*, pages 495-524, 2009.

[14] D. Csetverikov. Basic Algorithms for Digital Image Analysis. *Institute of Informatics, Budapest*, 2009.

[15] M.B. Ahmad and T.S. Choi. Local Threshold and Boolean Function Based Edge Detection. *IEEE Transactions on Consumer Electronics*, 45(3):674-679, August 1999.

[16] S. Wesolkowski and E. Jernigan. Color Edge Detection in RGB Using Jointly Euclidean Distance and Vector Angle. *In Proc. Of the IAPR Vision Interface Conference*, pages 19-21, May 1999.

[17] E. Nadernejad, S. Sharifzadeh and H. Hassanpour. Edge Detection Techniques: Evaluations and Comparisons. *Applied Mathematical Sciences*, 2(31):1507–1520, 2008.

[18] H. Deng, W. Zhang, E. Mortensen, T. Dietterich and L. Shapiro. Principal Curvature-Based Region Detector for Object Recognition. *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1-8, July 2007.

[19] K.G. Derpanis. The Harris Corner Detector. Technical report, York University, October 2004.

[20] F. Mokhtarian and R. Suomela. Robust Image Corner Detection Through Curvature Scale Space. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(12):1376–1381, December 1998.

[21] C. Schmid, R. Mohr and C. Bauckhage. Evaluation of Interest Point Detectors. *International Journal of Computer Vision*, 37(2):151–172, June 2000.

[22] C. Harris and M. Stephens. A Combined Corner and Edge Detector. *In Alvey Vision Conference*, pages 147-151, 1988.

[23] J. van der Weijer, T. Gevers and J.M. Geusebroek. Edge and Corner Detection by Photometric Quasi-Invariants. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(4): 625–630, April 2005.

[24] S. Ando. Image Field Categorization and Edge/Corner Detection from Gradient Covariance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(2):179–190, February 2000.

[25] D. Parks and J.P. Gravel. Corner Detection. URL *http://www.cim.mcgill.ca/dparks/CornerDetector/harris.ht*.

[26] D.G. Lowe. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, 60(2):91–110, January 2004.

[27] D.G. Lowe. Object Recognition from Local Scale-Invariant Features. *International Journal of Computer Vision*, 2:1150-1157, September 1999.

[28] D.G. Lowe. Local Feature View Clustering for 3D Object Recognition. *IEEE Conference on Computer Vision and Pattern Recognition*, 1:682-688, December 2001.

[29] M. Aly. Face Recognition using SIFT Features. Computer Science Department, California Institute of Technology, 2006.

[30] Z. Hua, Y. Li and J. Li. Image Stitch Algorithm Based on SIFT and MVSC. *Seventh International Conference on Fuzzy Systems and Knowledge Discovery*, pages 2628-2632, September 2010.

[31] Z. Chaoyang. Video Object Tracking using SIFT and Mean Shift. Master's thesis, Chalmers University of Technology, Sweden, 2011.

[32] L. Lei. Three Dimensional Shape Retrieval using Scale Invariant Feature Transform and Spatial Restrictions. Technical report, National Institute of Standards and Technology, August 2009.

[33] S. Se and P. Jasiobedzki. Stereo-Vision Based 3D Modeling and Localization for Unmanned Vehicles. *International Journal of Intelligent Control and Systems*, 13(1):46–57, March 2008.

[34] I. Pitas. Digital Image Processing Algorithms and Applications. John Wiley and Sons, Inc., 2000.