

Automatic Landmark Detection for Cervical Image Registration Validation

Juan D. García-Arteaga and Jan Kybic

Center for Machine Perception, Czech Technical University in Prague,
Karlovo náměstí 12, 121 35, Praha 2, Czech Republic

ABSTRACT

Many cervical Computer-Aided Diagnosis (CAD) methods rely on measuring gradual appearance changes on the cervix after the application of a contrast agent. Image registration has been used to ensure pixel correspondence to the same tissue location throughout the whole temporal sequence but, to date, there is no reliable mean of testing its accuracy to compensate for patient and tissue movement.

We present an independent system to use automatically extracted and matched features from a colposcopic image sequence in order to generate position landmarks. These landmarks may be used either to measure the accuracy of a registration method to align any pair of images from the colposcopic sequence or as a cue for registration. The algorithm selects sets of matched features that extend through the whole image sequence allowing to locate, in a reliable and unbiased way, a tissue point throughout the whole image sequence. Experiments on real colposcopy image sequences show that the approach is robust, reliable, and leads to geometrically coherent sets of landmarks that correspond to visually recognizable regions. We use the extracted landmarks to test the precision of some of the cervical registration algorithms previously presented in the literature.

Keywords: Registration, Computer-Aided Diagnosis, Optical, Colposcopy, Validation

1. INTRODUCTION

Uterine cervical cancer is the second most common form of cancer among women worldwide. Unlike other more aggressive cancers, cervical cancer tends to be slow growing, with a development rate of several years. During the early pre-cancerous stages it may be completely asymptomatic¹ with preinvasive cervix lesions detectable by specific screening methods. When detected and treated early, cervical cancer has a recovery rate of nearly 100%. This has led to an increasing demand for early detection methods with heightened sensitivity.

Colposcopy is a well established diagnostic method to detect cancerous and pre-cancerous tissue through visual inspection of the cervix.² During the exam the physician (colposcopist) applies a low concentration acetic acid (3-5%) on the cervix inducing color and texture changes in abnormal and metaplastic epithelia. The maximum contrast is achieved after 60 seconds followed by a slow return of the cervix to its original color after three to five minutes. The colposcopist observes the color changes through a low magnification microscope (colposcope), reports his findings and, if necessary, recommends a biopsy to confirm the diagnosis. Envisioned CAD diagnostic systems aim to aid the physician in his diagnosis by quantitatively measuring and combining features with high prognostic values by analyzing the images captured during the colposcopy exam.³ This will result in more reliable diagnosis, minimization of the variability among colposcopists and the reduction of unnecessary biopsies.

Because of patient and tissue movement, analysis involving multiple images (e.g. progressive texture or color change) are not possible without the previous registration of all images involved to a common frame. Several

Juan D. García-Arteaga: Email: garcia@fel.cvut.cz, Telephone: +420 224 355 734

Copyright 2007 Society of Photo-Optical Instrumentation Engineers. This paper was published in *Proceedings of SPIE, Volume 6514. Medical Imaging 2007: Computer-Aided Diagnosis* and is made available as an electronic reprint with permission of SPIE. One print or electronic copy may be made for personal use only. Systematic or multiple reproduction, distribution to multiple locations via electronic or other means, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper are prohibited

methods have been proposed to perform this registration for colposcopic images^{4,5,6,7} but, to the best of our knowledge, none of these methods was quantitatively validated. The main difficulty evaluating registration accuracy for cervical images is that presently there is no *gold standard* nor data sets with a known ground truth. Additionally, the tissue’s elasticity makes the true form of the spatial transformation model unknown. External markers such as small paper dots attached to the cervix or marks made with a surgical pen have been used for referencing purposes⁸ but their use as fiducial markers is limited as the paper dots may move and the ink diffuses.

It is possible to measure the accuracy of the registration by identifying the location of corresponding salient features in both the template and the un-registered image and then again in the registered colposcopic image.⁹ However, the lack of clear-cut anatomical landmarks makes it impossible for a human operator to consistently and reliably choose corresponding landmarks throughout the image sequence with the required precision. Furthermore, landmark choosing, even when done by an expert, is “subjective, difficult to reproduce and possibly erroneous”.¹⁰

As an alternative to this we have chosen to extract and match landmarks from the colposcopic image sequence by an automatic procedure. The algorithm produces a series of landmarks corresponding to the same tissue point throughout the whole sequence. These landmarks can then be used to evaluate the quality of the registration between any 2 images of the sequence.

The rest of the paper is organized as follows: Section 2 explains the landmark detection algorithm and presents some examples. Section 3 briefly reviews some of the registration methods used with cervical images which we have tested using the extracted landmarks and the data set used to test them. The results of these tests are discussed in Section 4. Conclusions and suggested future lines of research are presented in Section 5.

2. LANDMARK DETECTION

Our aim is to detect and correctly match corresponding points throughout all the images. These points might not always correspond to anatomical landmarks, but rather to locations that can be clearly identified in any of the sequence’s images.

The process can be summarized in four steps: Sequential pairwise alignment, feature extraction, feature matching and landmark selection. We will detail each of these steps in the rest of this section.

2.1. Sequential Pairwise Alignment

Because the displacement between images may be large we have opted to first roughly align them by a rigid transformation. This, while leaving the local appearance mostly unchanged, reduces significantly the search radius and the number of mismatches in the subsequent matching step and accelerates the algorithm.

Rigid (or orthonormal) transformations are composed of rotations and translations. They preserve distances and angles between point locations and, in homogenous coordinates, can be represented by a 3x3 matrix H . Their inverse transformation is easily calculated by inverting matrix H .

One option is to align all frames to one common template frame (for example the first). However, during the colposcopic examination the appearance of the cervix varies with time because of the gradual color changes induced by the acetic acid and other sources such as illumination differences and the appearance of new objects in the field of view (e.g. glints, mucous, blood spots). This makes the registration between any two frames of a colposcopic sequence increasingly difficult with increasing time intervals between their capture. Instead, we apply the pairwise sequential registration scheme shown in Figure 1.

In this scheme the second frame (Img2) is registered to the first frame (Img1) and the resulting image (Img2’) and the transformation ($H_{1 \rightarrow 2}$) are saved. The scheme is iterated registering Img(n) to Img(n-1)’ until all images are registered. This will compensate for most of the gross distortion due to camera movement. For the pairwise registration we have used the algorithm developed and implemented by Thévenaz *et al*¹¹ because of its robustness.

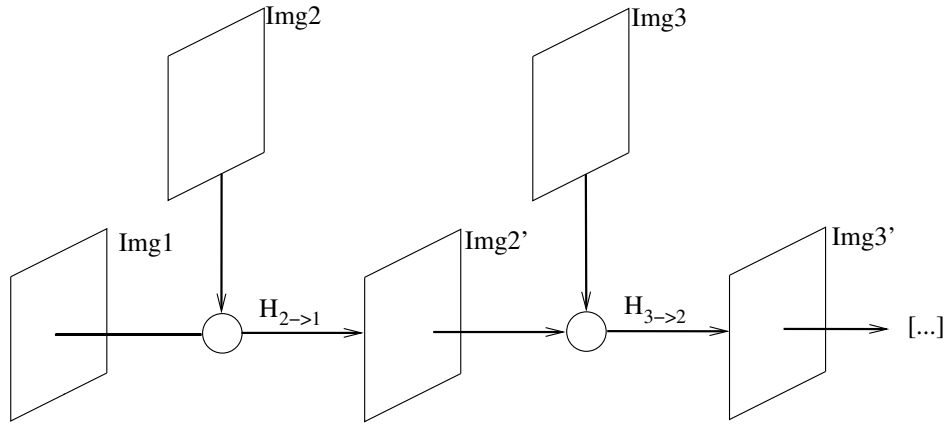


Figure 1. Schematic representation of the registration scheme used for the initial alignment of the frames. Each circle represents the pairwise rigid registration algorithm.

2.2. Feature Extraction

Feature points that serve as landmark candidates are extracted using the Harris detector¹². This well studied method detects sharp intensity changes in the first derivative of an image, reliably detecting features such as corner points and, to a lesser degree, tight curves. Its rotational invariancy and partial invariance to affine intensity changes make it reliable and reproducible. It has been shown previously⁴ that the center of the cervical region has textures and features with borders sharp enough to be found by the Harris detector. An example of the interest points extracted from two consecutive frames of a high resolution colposcopy sequence is shown in Figure 2.

2.3. Feature Matching

As color and textural changes are gradual, for images captured at short time intervals good putative matches can be identified between the features found by the Harris detector. Matches are made between all consecutive frames: any two feature points in consecutive images are matched if they are maximally correlated to each other within a neighboring window and if the distance between them is less than a predefined search radius. Peter Kovesi's Matlab implementation¹³ of the normalized cross-correlation¹⁴ was used over a 7x7 window.

An example of putative matches between consecutive registered and unregistered frames may be seen in Figure 2. The search radius was set to 250 and 15 pixels for the registered and unregistered images, respectively.

2.4. Landmark Pruning

To select the candidate landmarks between images m and n of the sequence, matches are made between all consecutive sequence frames. Only the points of interest that are matched in all frames between m and n are chosen, that is, for the candidate feature point j in the initial image m (\mathbf{X}_m^j) and the candidate feature point j in the last image n (\mathbf{X}_n^j) to be accepted there must exist an equivalent candidate feature point in image k for every k between m and n (\mathbf{X}_k^j). In other words, $(\mathbf{X}_m^j, \mathbf{X}_n^j)$ is a candidate pair if and only if:

$$\forall k, k \in \mathbb{N}, m \leq k < n \quad \exists \quad \mathbf{X}_k^j, \mathbf{X}_{k+1}^j; \mathbf{X}_k^j \leftrightarrow \mathbf{X}_{k+1}^j \quad (1)$$

where $\mathbf{Y} \leftrightarrow \mathbf{Z}$ represents a putative match between points \mathbf{Y} and \mathbf{Z} . The logic behind this is that weak or rapidly changing features (such as glint) will not last for many frames. This reduces significantly the number of mismatches. As a final step, a human operator can check the results and reject any remaining mismatches. Some of the landmarks chosen in different cervical sequences may be seen in Figure 3.

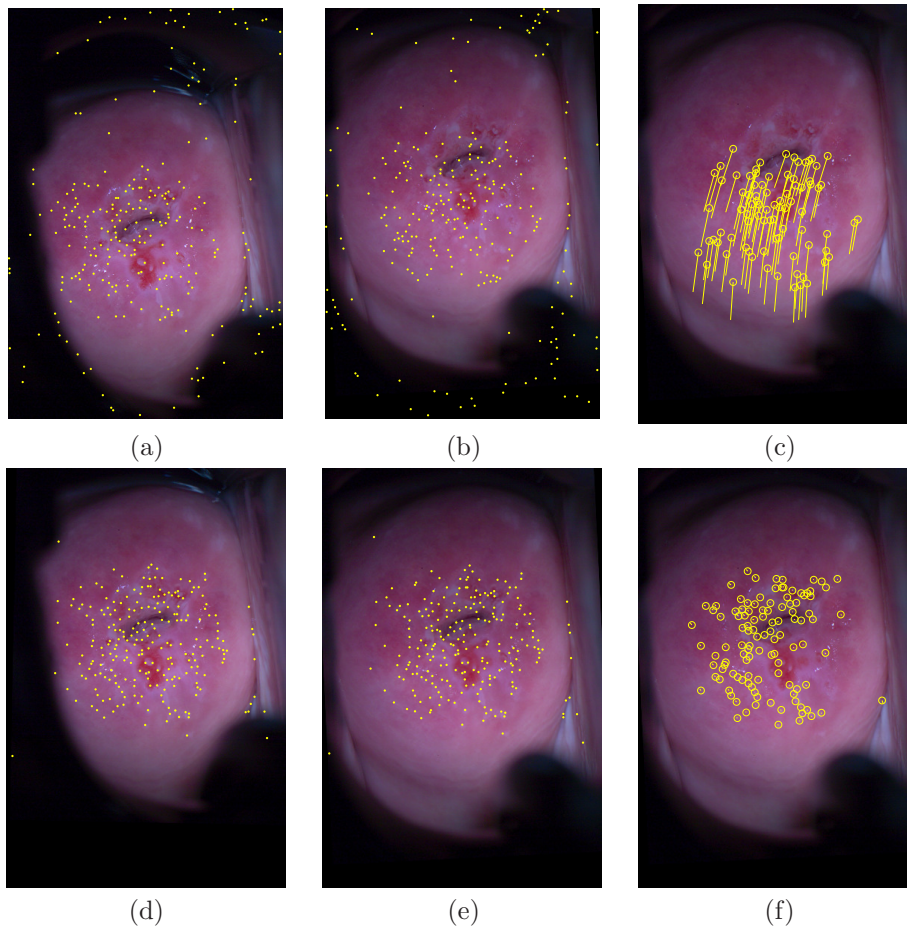


Figure 2. Two cervical images taken at 60 and 73 seconds after the application of acetic acid with the detected features superimposed (images (a) and (b)). Correct matches for the detected features are shown as lines connecting the correctly matched features' position on the 73 seconds image (image (c), circles at the end of the line show the features detected in the displayed image). Results of the same procedure after rigidly registering the images are shown on the second row (images (d), (e) and (f)).

2.5. Landmark Validation

Visual inspection is, presently, the final landmark validation; it is a quick and simple approach that has been proven to accurately detect misregistrations in other image modalities: for MRI to CT Brain images 2-millimeter misregistrations could be detected and in PET-MRI images 2-millimeter displacements along the x and y axes⁹. It should be noted that this degree of accuracy was achieved on dissimilar/multimodal images, whereas we deal with a set of highly detailed images, all captured by the same device. Although no similar study has been conducted for cervical images the consistency of the landmarks in all images of the sequence (as seen on Figure 3) leads us to believe that visual inspection is sufficient.

3. EXPERIMENTS

Landmarks extracted with the method described in Section 2 were used to test the precision of four registration algorithms previously proposed for colposcopy images. The matched landmarks used may be seen overlaid on the initial and final image of the sequences on Figure 4.

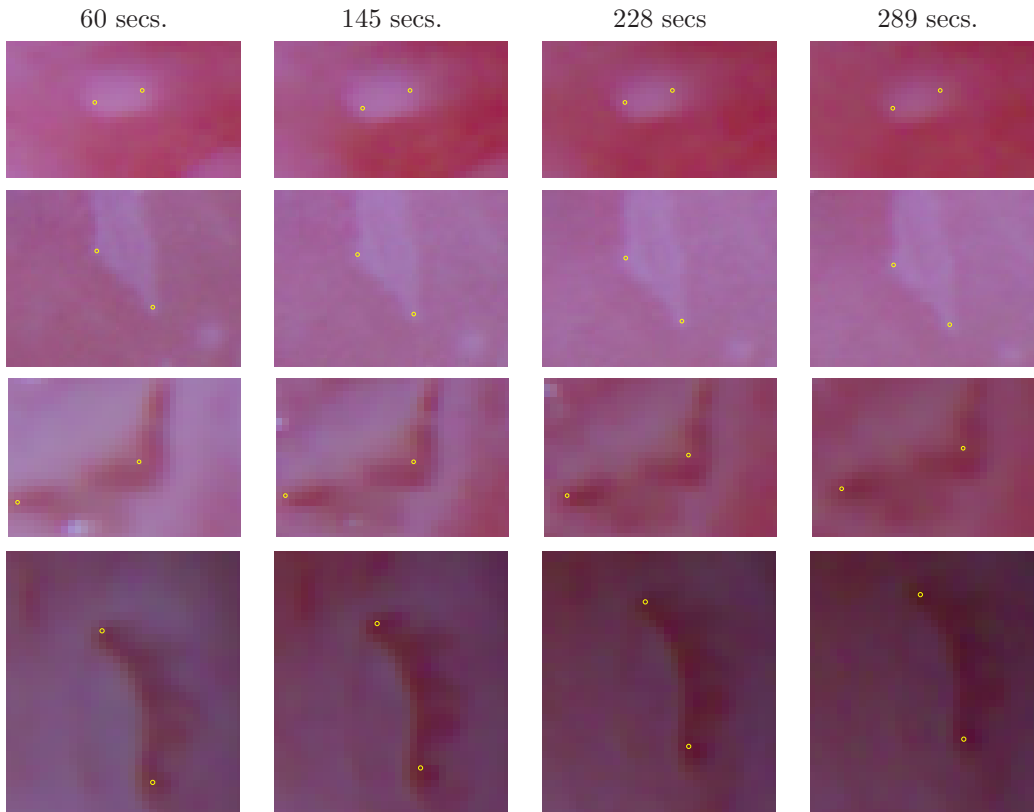


Figure 3. Corresponding features over time. The features were cropped from the images taken at, from left to right: 60, 145, 229 and 289 seconds after acetic acid application .

3.1. Tested Methods

There is an increasing amount of medical and academic literature documenting the advantages of digital colposcopy as a screening and diagnosis tool for uterine cervical cancer^{3,15} . Methods that compare multiple images must at some point use image registration to bring all analyzed images to a common frame. Rigid displacement based on phase correlation over a network of overlapping windows was used to analyze temporal changes in colposcopic video images⁶ . A stabilization algorithm based on the composition of the homographies of subsequent frames and heuristics was used to analyze temporal texture changes in cervical video sequences⁴ . An elastic image registration based on B-splines¹⁶ previously used in Echo Planar Imaging (EPI) and Magnetic Resonance Imaging (MRI)¹⁷ was used to fuse hyperspectral data from a cervical imaging instrument⁵ . Elastic image registration using optimization over a set of continuous deformation vector fields was used for movement compensation in high resolution colposcopy images⁷ .

We have tested these four registration methods (**PHASE**⁶ , **HOMOGRAPHY**⁴ , **BSPLINE**⁵ and **VECTOR**⁷)with real colposcopic images. The landmarks used as ground truth for these images were extracted by the procedure explained in Section 2. Except for **HOMOGRAPHY**, that finds the transformation between two images by composing the transformations of all intermediate frames, all other methods use only the reference and a test image images to calculate the deformation between them.

Two of the tested methods, **VECTOR** and **SPLINES**, register the images by minimizing a dissimilarity measure between them (in both cases the Sum of Square Differences or SSD). As they both rely on local iterative optimization algorithms the final solution will depend on the starting point. As seen on Figure 2, sometimes there is a severe camera movements between consecutive frames. This leaves the iterative methods too far away from the global optimum to converge to the correct answer. To compensate for this we run two additional tests

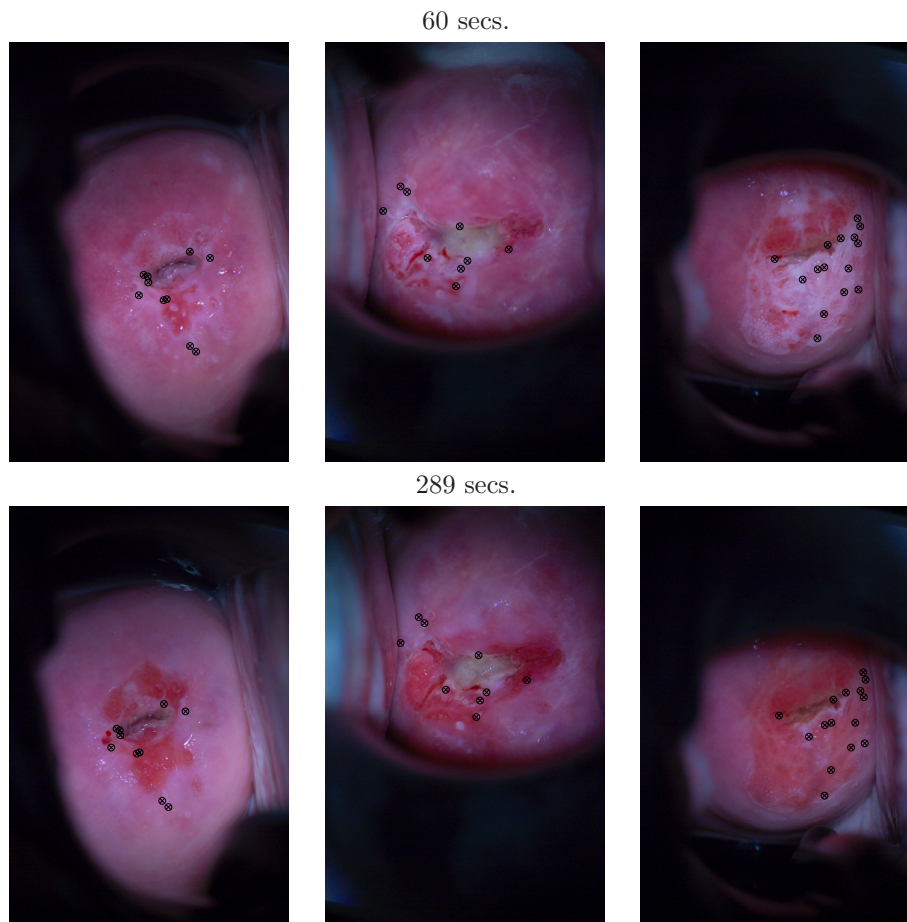


Figure 4. The final landmarks are overlaid on the initial and final frame of the sequence (60 and 299 seconds after acetic acid application, top and bottom rows respectively). From left to right sequences 82, 90 and 98.

(VECTOR2 and SPLINES2) using the results of the sequential rigid registration explained in Section 2.1 as the starting point.

3.2. Test Images

Three sequences composed of 20 images taken at intervals of between 5 and 12 seconds during colposcopic exams of different patients were used to test the methods explained in Section 3.1. The test image was captured 60 seconds after the application of acetic acid (moment of maximal acetic reaction, when the cervix is whitest) and corresponds to the initial frame. The reference image is the final frame, captured 289 seconds after the application of the acetic solution, when the cervix has almost completely returned to its normal coloration.

The images are originally RGB 16-bit 750x1125 pixels taken with a cross-polarized filter to reduce glint and with a black opaque speculum to cancel the background. All methods convert the images to grayscale.

3.3. Region of Interest

Four concentric areas of the cervix can be seen on the images: 1) The Os region, at the center, is the dark elongated or circular zone that corresponds to the end of the uterine cavity, 2) the Columnar Epithelium (CE), surrounding the Os, which appears red and irregular, 3) The Squamous Epithelium (SE), which appears as a smooth, pink featureless tissue, and 4) the Transformation Zone (TZ) which lies between the CE and the SE. The vagina walls and the speculum used to separate them may be seen also. The most interesting zone for

diagnostic purposes, and therefore the area on which the registration should concentrate, are the CE and the TZ. To reflect this, region of interest masks (ROIs) containing, approximately, the CE and TZ was chosen and used during the registration process. Alternatively, some proposed cervical segmentation methods^{18,19,20} may be used to automate the ROI selection process.

The implementation of **SPLINE** and **SPLINE2** we have used* does not allow the use of ROI masks. Because of this, no ROI was used in **SPLINE** and a cropped version of the rigidly registered images containing the ROIs was used for **SPLINE2**.

3.4. Results

All images of the sequence are registered to the reference frame. The registration error for each image pair is calculated as the average of the distances between the landmark position in the reference image and the landmark position in the registered test image:

$$Err = \frac{1}{N} \sum_{j=1}^N |H(\mathbf{X}_{test}^j) - \mathbf{X}_{ref}^j| \quad (2)$$

$$H(\mathbf{X}_{test}^j) = \mathbf{X}_{registered}^j \quad (3)$$

where \mathbf{X}_{test}^j and \mathbf{X}_{ref}^j are the coordinates of the j th landmark pair and H is the transformation found by the registration. The graphs of Error vs. Capture time difference for each of the tested methods may be seen in Table 1.

4. DISCUSSION

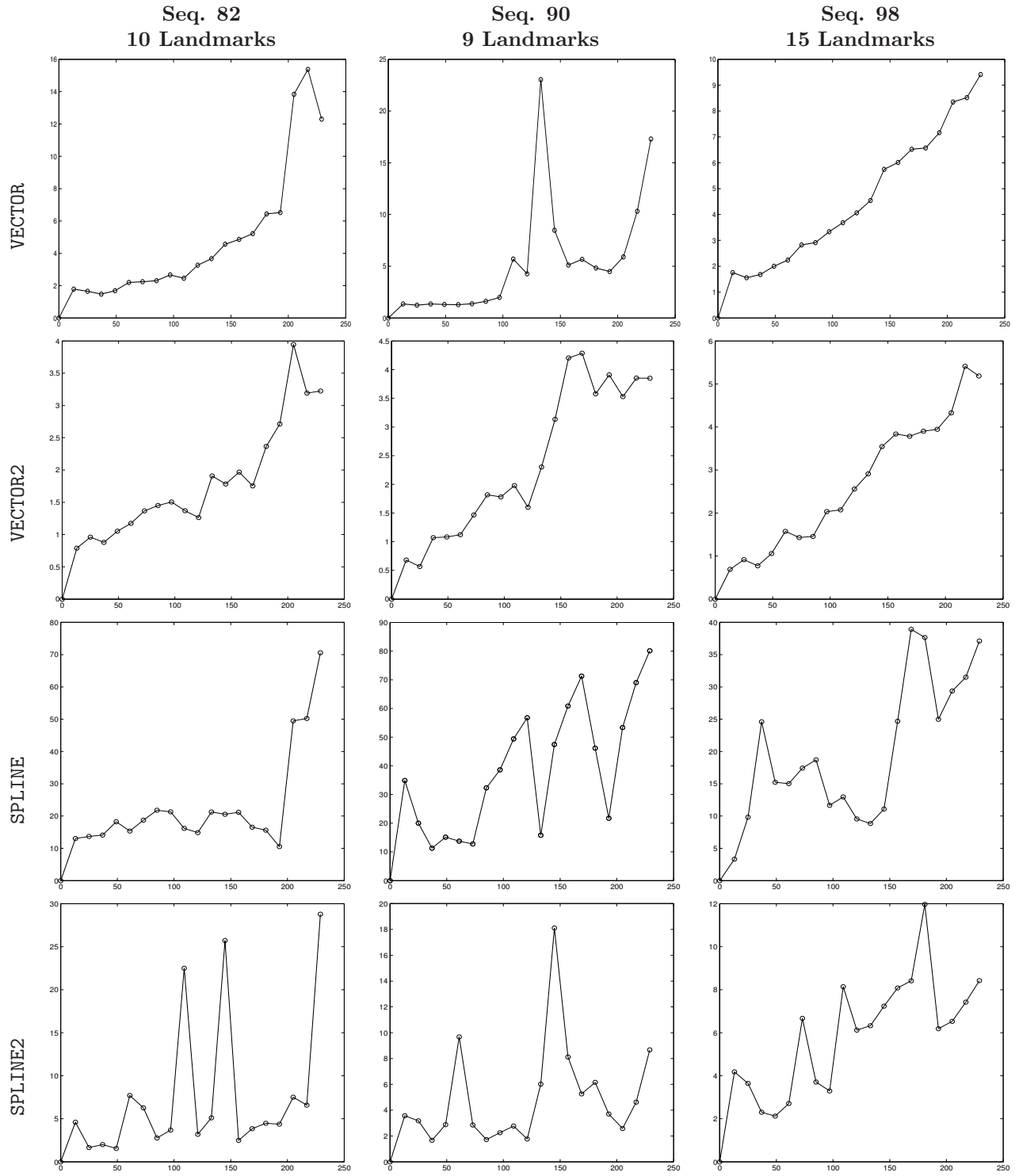
As expected, all methods perform better for images captured at short intervals. The progressive loss of precision as capture time difference between the images grows may have a number of causes: Errors have accumulated (for algorithms that use intermediate results), the deformation model is no longer able to correctly describe the image changes or the similarity measure used is no longer valid. It is difficult to measure the relative contribution of each of these factors in the total error; however, the good overall performance of the **HOMOGRAPHY** method with a relatively simple deformation model (8 degrees of freedom) indicates that, although present, the elastic movement of the tissue doesn't seem to be the predominant source of image misalignment. This would also explain the good performance of **VECTOR2** by using the results of the rigid registration scheme presented in section 2.1 as a starting point.

As noted before, all methods based on iterative optimization (**SPLINE**, **SPLINE2**, **VECTOR** and **VECTOR2**) used SSD as a similarity criterion. This makes them particularly sensitive to the color and illumination changes. This is more noticeable on the **SPLINE** method, as it allows large soft deformations while looking for the minimum. The SSD criterion, although simple and fast, might not be enough to measure the similarity of the images. More complex similarity criteria, such as the ones used in multi-modal image registration, may significantly improve the results of these methods.

The **PHASE** method was found to be useful only for very low resolution images. The images on which it was originally tested were processed had a resolution of 74x99,⁶ which is roughly 10% of the one used for the present tests.

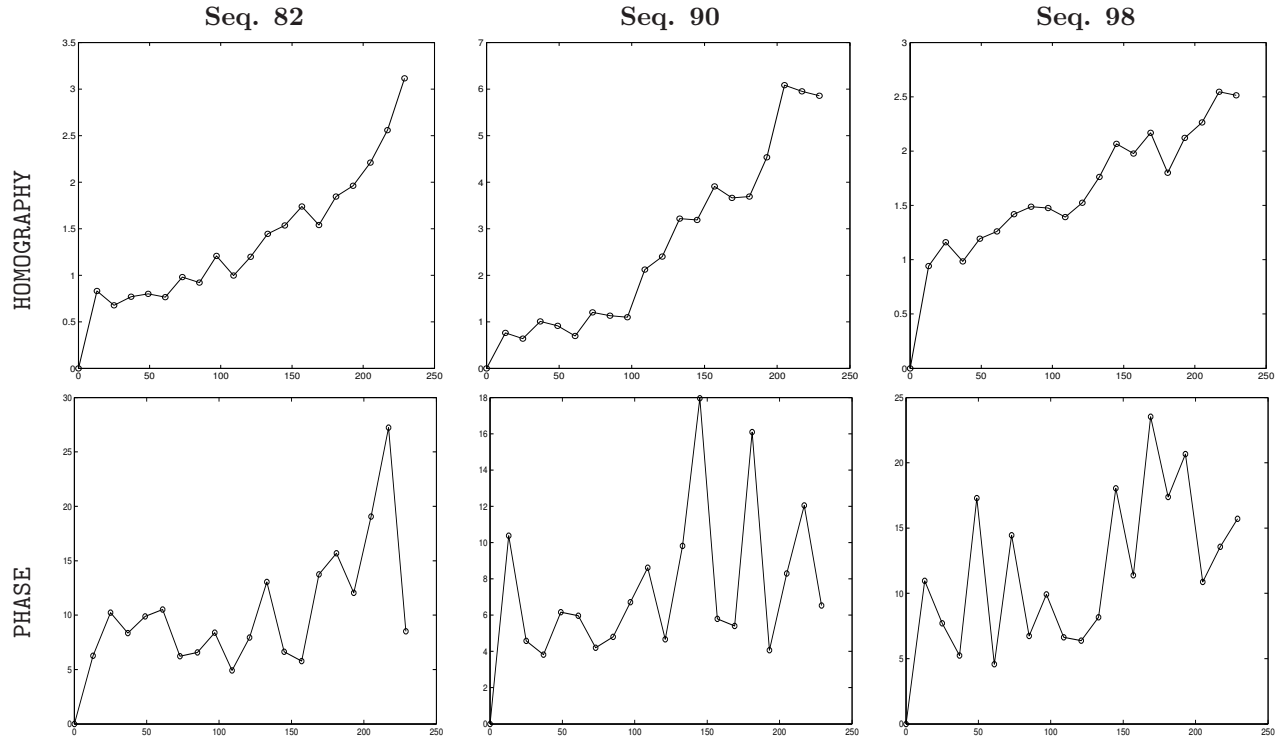
*Available from the author's web site: <http://cmp.felk.cvut.cz/~kybic/thesis/>

Table 1: Average Error vs. Capture Time Difference.



Continued on next page

Table 1 – continued from previous page



5. CONCLUSIONS AND FUTURE WORK

A method to extract landmarks from colposcopic image sequences was presented. The extracted landmarks were used to evaluate the precision of some of the registration algorithms used for colposcopic images and to compare their performance on images acquired during a typical exam.

The system was able to extract landmarks corresponding to visually recognizable regions, thus avoiding the bias present in human landmark selection. These landmarks may be used either to validate future registration algorithms or as a valuable cue to guide the image registration process itself.

The test results indicate that the use of low complexity deformations, improved similarity criteria and a better integration of the intermediate results are the keys to a better performance in cervical image registration.

Finally, it should be noticed that digital colposcopy is a relatively new method on which *de facto* image quality standards have not yet emerged. This leads researchers from different groups to test their methods on images of varying qualities and characteristics. We have tried to overcome this and make a fair comparison. Future tests will be able to pinpoint more precisely the advantages and disadvantages of each method, independently of the dataset, as the image quality becomes more standardized.

6. ACKNOWLEDGEMENTS

The authors were supported by The Grant Agency of the Czech Academy of Sciences under project 1ET101050403.

REFERENCES

1. T. P. Canavan and N. R. Doshi, "Cervical cancer," *American Family Physician* **61**(5), pp. 1369–1376, 2000.
2. D. Ferris, *Modern Colposcopy: Textbook and Atlas*, Kendall Hunt Publishing Company, Dubuque, Iowa, 2nd ed., 2002.

3. H. Lange and D. G. Ferris, "Computer-aided-diagnosis (CAD) for colposcopy," *Medical Imaging 2005: Image Processing* **5747**(1), pp. 71–84, 2005.
4. A. Barreto-Flores, L. Altamirano-Robles, R. M. Morales-Tepalt, and J. D. Cisneros-Aragon, "Identifying precursory cancer lesions using temporal texture analysis.," in *Second Canadian Conference on Computer and Robot Vision*, pp. 34–39, IEEE Computer Society, 2005.
5. H. Lange, R. Baker, J. Håkansson, and U. P. Gustafsson, "Reflectance and fluorescence hyperspectral elastic image registration," in *Medical Imaging 2004: Physiology, Function, and Structure from Medical Images. Edited by Amini, Amir A.; Manduca, Armando. Proceedings of the SPIE, Volume 5370, pp. 335-345 (2004).*, J. M. Fitzpatrick and M. Sonka, eds., pp. 335–345, May 2004.
6. H. G. Acosta-Mesa, H. V. Ros-Figueroa, N. Cruz-Ramírez, A. Marín-Hernández, B. Zitová, R. Hernandez-Jimenez, B. E. Cocotle-Ronzon, and E. Hernandez-Galicia, "Cervical cancer detection using colposcopic images: a temporal approach.," in *Sixth Mexican International Conference on Computer Science*, pp. 158–164, 2005.
7. J. D. García-Arteaga, J. Kybic, and W. Li, "Elastic image registration for movement compensation in digital colposcopy," in *Analysis of Biomedical Signals and Images. BIOSIGNAL 2006*, J. Jan, J. Kozumplík, and I. Provazník, eds., pp. 236–238, Brno University of Technology, June 2006.
8. D. G. Ferris, R. A. Lawhead, E. D. Dickman, N. Holtzapple, J. A. Miller, S. Grogan, S. Bambot, A. Agrawal, and M. L. Faupel, "Multimodal hyperspectral imaging for the noninvasive diagnosis of cervical neoplasia," *Journal of Lower Genital Tract Disease* **5**(2), pp. 65–72, 2001.
9. R. P. Woods, "Validation of registration accuracy," in *Handbook of medical imaging*, pp. 491–497, Academic Press, Inc., (Orlando, FL, USA), 2000.
10. P. Rogelj, S. Kovacic, and J. Gee, "Validation of a nonrigid registration algorithm for multimodal data," in *Proceedings of SPIE Medical Imaging 2002, Image Processing*, J. M. Fitzpatrick and M. Sonka, eds., **4684**, pp. 299–307, SPIE, February 2002.
11. P. Thévenaz, U. Ruttimann, and M. Unser, "A pyramid approach to subpixel registration based on intensity," *IEEE Transactions on Image Processing* **7**, pp. 27–41, January 1998.
12. C. Harris and M. Stephens, "A combined corner and edge detection," in *Proceedings of The Fourth Alvey Vision Conference*, pp. 147–151, 1988.
13. P. D. Kovesi, "MATLAB and Octave functions for computer vision and image processing." School of Computer Science & Software Engineering, The University of Western Australia. Available from: <http://www.csse.uwa.edu.au/~pk/research/matlabfns/>.
14. L. G. Brown, "A survey of image registration techniques," *ACM Comput. Surv.* **24**(4), pp. 325–376, 1992.
15. D. G. Ferris, "Analysis of digitized cervical images to detect cervical neoplasia," in *Medical Imaging 2004: Image Processing. Edited by Fitzpatrick, J. Michael; Sonka, Milan. Proceedings of the SPIE, Volume 5370, pp. 181-194 (2004).*, J. M. Fitzpatrick and M. Sonka, eds., pp. 181–194, May 2004.
16. J. Kybic, *Elastic Image Registration Using Parametric Deformation Models*. EPFL thesis no. 2439 (2001), 180 p., Swiss Federal Institute of Technology Lausanne (EPFL), July 26, 2001.
17. J. Kybic and M. Unser, "Fast parametric elastic image registration," *IEEE Transactions on Image Processing* **12**, pp. 1427–1442, November 2003.
18. S. Gordon, G. Zimmerman, and H. Greenspan, "Image segmentation of uterine cervix images for indexing in pacs.," in *CBMS*,²¹ pp. 298–303.
19. H. Lange, "Automatic detection of multi-level acetowhite regions in RGB color images of the uterine cervix," in *Medical Imaging 2005: Image Processing. Edited by Fitzpatrick, J. Michael; Reinhardt, Joseph M. Proceedings of the SPIE, Volume 5747, pp. 1004-1017 (2005).*, J. M. Fitzpatrick and J. M. Reinhardt, eds., pp. 1004–1017, Apr. 2005.
20. P. King, S. Mitra, and B. Nutter, "An automated, segmentation-based, rigid registration system for cervigramtm images utilizing simple clustering and active contour techniques.," in *CBMS*,²¹ pp. 292–297.
21. *17th IEEE Symposium on Computer-Based Medical Systems (CBMS 2004), 24-25 June 2004, Bethesda, MD, USA*, IEEE Computer Society, 2004.