International Journal of Computational Vision and Biomechanics
Vol. 3 No. 1 (January-June, 2017)

MUK PUBLICATIONS
Open Access Publisher

Automated Overlay of Infrared and Visual Medical Images using Skin Detection and Image Registration

G. Schaefer[1], R. Tait[2], A. Merla[3] and S.Y. Zhu[4]

[1]*Department of Computer Science, Loughborough University, U.K.*

[2]*Department of Experimental Psychology, University of Cambridge, U.K.*

[3]*Department of Clinical Sciences and Biomedical Imaging, University of Chieti-Pescara and Institute for Advanced Biomedical Technology, Foundation G. D'Annunzio University, Italy*

[4]*School of Computing, University of Derby, U.K.*

Thermal infrared imaging is used to capture the temperature distribution of the human skin and employed in various medical applications. Often visual images are taken in conjunction with thermograms. Clinicians are interested to cross-reference the infrared and visual images of a patient, either to see which part of the anatomy is affected by a certain disease or to judge the efficacy of treatment. In this paper, we show that image registration techniques can be used effectively to generate such an overlay of visual and thermal infrared images to provide a useful visualisation. Following a skin detection step, which segments areas corresponding to the patient in the visual image, registration techniques, either intensity- or landmark-based, are employed to accurately align the two images. The proposed approach is shown to be effective for a variety of applications, ranging from clinical studies such as monitoring the evolution of lesions in dermatology (psoriasis, dermatitis) or in immunology (scleroderma, lupus), to the most innovative applications of functional infrared imaging, like emotion recognition through a combination of infrared imaging-based computation of autonomic responses and facial expression recognition.

1. INTRODUCTION

Thermal infrared imaging is a non-invasive, non-contact, passive, and radiation-free imaging modality to capture the natural thermal radiation generated by an object at a temperature above absolute zero. Often, visual images are captured in conjunction with thermograms, for example to relate inflamed skin areas to the human anatomy, which in turn is useful for medical diagnosis as well as for assessing the efficacy of any treatment. Currently, this process requires great expertise and is subject to the individual clinician's ability to mentally map the two distinctly different images.

Image registration is the process of geometrically aligning or overlaying two images taken from different sensors or perspectives and/or at different times, and is a crucial part of many medical imaging applications. Images are aligned by solving for the optimal transformation, expressed as a combination of scaling, translation and rotation, that will map information from one image to the other one. Registration is often used to monitor growth, verify the effects of treatment or to make comparisons of patient data with anatomically normal subjects.

In this paper, we show how image registration can be effectively used to overlay thermal infrared images and visual images of a patient in order to relate areas that are of interest due to their thermal pattern to the human anatomy. After capturing both infrared and visual images, the visual image undergoes a skin detection step that is used to separate the patient from the background. Intensity-based image registration, which requires no user interaction, is then employed to superimposed the two images and the generated overlay is presented to the user. Should this automatic approach lead to inaccurate results, in a further step a landmark-based registration technique is utilised where alignment is based on control points defined by the user. The generated system was employed at the Institute for Advanced Biomedical Technology at Foundation G. D'Annunzio University in Chieti, for, among others, studying the evolution of the clinical scenario in patients suffering for Raynaud's phenomenon and scleroderma [12, 15], and analysis of the emotional response in individuals exposed to audio-video stimuli through autonomic thermal effects and facial expressions [18, 14].

2. BACKGROUND

2.1. Thermal Infrared Imaging

Thermal infrared imaging uses a camera with sensitivities in the infrared to provide a picture of the temperature distribution of human skin [7]. It is a non-invasive, non-contact, passive, radiation-free technique that can be employed in combination with anatomical investigations (e.g., based on x-rays or CT/MRI investigations) and can

reveal problems when the anatomy is otherwise normal. The radiance from human skin can be characterised as an exponential function of the surface temperature which in turn is influenced by the level of blood perfusion in the skin. Thermal infrared imaging is therefore well suited to pick up changes in blood perfusion which might occur due to inflammation, angiogenesis or other causes. It has been shown early [21] that asymmetries in temperature distributions as well as hot and cold spots are strong indicators of an underlying dysfunction [13]. Thermal infrared imaging has been successfully employed to detect breast cancer [1, 5], diagnose Raynaud's phenomenon [11], and to monitor local scleroderma [2]. Various image processing and analysis techniques have been successfully utilised in acquiring and interpreting medical infrared images (see e.g. [17, 23]) and have been shown to be useful and important tools for clinical diagnosis.

2.2. Image Registration

Image registration is used to geometrically align two images taken from different sensors or perspectives and/ or different times. One of the images is assigned as reference (or fixed) image while the other serves as sensed (or moving) image. The sensed images is then transformed using a combination of scaling, translation and rotation operations so as to best match the reference image. However, finding this best fit is intrinsically difficult because of the various geometric deformations caused by the diverse methods of image capture as well as by image noise and other factors. In medical image analysis, registration is often used to monitor growth, verify the effects of treatment or to make comparisons of patient data with anatomically normal subjects [10].

Image registration is an intensive research area with many registration techniques introduced in the literature. In general, most of these can be divided into two classes: landmark-based and intensity-based registration algorithms [24]. Landmark-based methods require specification of pairs of corresponding control points in both images. These control points are then utilised to determine the mapping between the two images where the goodness of fit is usually determined in terms of the deviations of the original and transformed landmarks. Obviously, in this approach the registration accuracy is inherently dependent on the accurate placement of the landmarks. Although careful manual specification of control points will typically lead to good registration accuracy, this method is time consuming and success relies heavily on the user's expertise. Control points can also be extracted automatically, however this proves difficult and unreliable in many cases and often requires certain domain specific knowledge. In contrast, intensitybased image registration is not based on control points. Here, the complete image data is used to arrive at an appropriate transform. This is achieved by iterating through transform optimisation, image resampling and feature-matching stages. Of these, feature matching is the most important part as it specifies a similarity metric which measures the fit between reference and sensed image.

3. METHODS

Our system is designed to provide an overlay of visual and thermal infrared images to be presented to the user/ clinician. Superimposing both image types allows to relate areas of certain thermal patterns to the anatomy of the person as well as to monitor efficacy of any treatment. After pre-processing, which importantly involves segmenting the image area corresponding to the patient through a skin detection step, intensity-based image registration based on a mutual information similarity metric is employed to align the two images. This is performed using a series of intelligent agents which collaborate on a blackboard architecture, providing an efficient and effective framework for image registration. Superimposed images are presented to the user for visualisation, where the user is able to adjust the relative importance of the individual modalities in an interactive manner.

3.1. Image Pre-processing and Segmentation

In our approach we are primarily employing intensity-based image registration to produce the overlay of visual and infrared images of patients. As mentioned above, intensity-based registration is able to utilise all available image data in order to arrive at an accurate mapping between the two images. However, this also means that any confounding image information that might lead to incorrect alignment needs to be dealt with. In particular, while the background area of infrared images taken in a temperature controlled lab is typically well separated from the patient's body, the same contrast must be achieved for the visual image. Whereas for infrared images a simple thresholding algorithm [16] is sufficient to segment the foreground (i.e. the patient) from the background, visual images need a more elaborate analysis. In our approach, we make use of the fact that infrared imaging picks up the skin temperature, and therefore employ a skin detection technique on the visual image to segment the area corresponding to the patient. For this, we adopt, with some variations, the method introduced in [4] which is based on the fact that the hues of human skin occupy only a small region in colour space. The algorithm, which operates on the visual (RGB) image, proceeds in the following steps:

Step 1: The R, G, and B values at each pixel are transformed into a log-opponent colour representation by

$$\begin{pmatrix} I \\ R_g \\ B_y \end{pmatrix} = \begin{pmatrix} L(G) \\ L(R) - L(G) \\ L(B) - \dfrac{L(G) - L(R)}{2} \end{pmatrix} \quad (1)$$

with

$L(C) = 105 \log_{10} (C + 1 + n)$ $C = \{R, G, B\}$

where n represents some random noise (in the range (0; 1)) to prevent banding artefacts in dark regions. I is the resulting intensity channel, whereas R_g and B_y are red-green and blue-yellow chromaticity channels.

Step 2: A measure of texture amplitude, T, is derived from the intensity channel I by calculating a difference image between the original image and a median filtered version of it, where the size of the median filter depends on the image resolution. The resulting texture channel T is then again median filtered as are the opponent channels R_g and B_y (at a finer scale compared to the texture channel).

Step 3: Hue H and saturation S are calculated as

$$H = \tan^{-1}\left(\frac{R_g}{B_y}\right)$$

$$S = \sqrt{R_g^2 + B_y^2} \quad (3)$$

Step 4: Pixels that fall within a certain hue-saturation range and do not exceed a texture threshold are identified. In particular, all pixels that fall within

$\{T < 5$ and $110 < H < 155$ and $5 < S < 60\}$

or

$\{T < 5$ and $130 < H < 170$ and $30 < S < 130\}$

are marked as skin pixels.

Step 5: Using morphological operations holes are filled and edges smoothed to provide the final output of the skin detector.

Step 6: Pixels that are not identified as part of skin regions are set to black, thus effectively removing the background.

3.2. Registration

After pre-processing, both the infrared and visual image will contain of a clearly separated foreground area corresponding to the patient, while the background is set to zero intensity in both images. Since, as mentioned above, landmark-based registration approaches are time-consuming, we are primarily interested in employing an intensity-based technique in order to provide a fully automatic method of generating visual-infrared overlays. In intensity-based methods, the choice of similarity metric is crucial for successful registration. Since the images are of different modalities, direct comparison of image intensities will lead to poor results, and we are therefore using a similarity metric based on the mutual information $MI(I, V)$ between infrared image I and visual image V, defined as [22]

$$MI(I, V) = H(V) - H(V \setminus I) \quad (4)$$

where $H(V)$ is the marginal entropy of the visual image while $H(V \setminus I)$ represents the joint entropy calculated from intensities of the infrared and visual images.

Not surprisingly, the similarity computation stage is complex and therefore represents a considerable performance bottleneck when employed in an iterative registration procedure. Based on a worker/manager model, we therefore employ a distributed blackboard system, where the workload is spread between a number of intelligent agents [20].

In this framework, the registration process begins with the partitioning of visual and thermal images into segments. After calculation of transform parameters, for each sample point in the visual segment a corresponding intensity in the infrared segment is calculated. Intensities at non-grid positions are derived using B-spline based interpolation [8]. A local Parzen histogram is then generated from retrieved intensities by a worker agent and placed onto the blackboard.

When all worker agents have terminated, the manager agent is able to construct a global Parzen histogram from the local histograms now available on the blackboard. By estimating the density distribution of the global histogram, an entropy value in the form of a gradient is then calculated. Regular step gradient descent optimisation [9] is employed by the manager agent in order to advance transform parameters in the direction of the gradient. During each iteration of the optimisation process, the step length through the transform search space is calculated using a bipartition scheme. Once updated, transform parameters are then propagated to all worker agents via the blackboard and the procedure is repeated until the algorithm has converged. Finally, the image segments are accumulated by the manager agent to assemble the registered sensed image.

Our employed framework is flexible and can be employed to either distribute the registration of single images as outlined above or to distribute the processing of many images in batch mode. It should also be noted that the framework can also be employed for registering 3-D volume datasets in an equally efficient and effective way [19].

After successful registration, a composite image is created. This is performed by computing the weighted

sum of the respective pixel values of the original visual image and the thresholded thermogram. Equal weights will generate an average of the two images whereas different weights will put more emphasis on one of the two modalities. The actual weightings between the two modalities can be determined by the user and can be controlled interactively.

Although the system by default performs intensity-based registration to provide a fully automated system, in some cases this approach does not produce an accurate overlay. In such instances the user has the possibility to switch to a landmark-based algorithm. Corresponding control points then need to be specified by the user in both the visual and the infrared image. Based on these landmarks the best transform is then calculated to align the two images.

4. RESULTS

Thermal images for the present study were acquired by means of a digital thermal camera (FLIR SC3000, FlirSystems, Sweden), with a focal plane array of 320x240 QWIP detectors, capable of collecting the thermal radiation in the 8-9 μm band, with a 0.02 second time resolution, and 0.02 K temperature sensitivity.

Figures 1 to 3 give examples of visual-infrared image pairs to show how our proposed method works. Each figure shows the original visual image, the thermal infrared image, the visual image segmented based on the output of the skin detection step, and the final, overlaid image. As can be seen, in all cases an accurate overlay of the two image types is achieved. It can also be noticed that the skin detection step does not have to perform perfectly in extracting the area corresponding to the patient (e.g., in Figure 1 areas corresponding to the hair and beard are (correctly) not classified as skin regions).

Figure 1 shows the distribution of the facial cutaneous temperature in a healthy subject. The generated images permit to identify the anatomical correspondence of the facial region exhibiting the largest temperature variations associated with emotional responses and expressions. This capability is very important as it permits to understand which association is set among facial (i.e., visual) and thermal (i.e, autonomic) expression of emotions. The output of Figure 1 allows for an enhanced Facial Action Coding System, which is at the basis of automatic recognition of facial expression [3].

Figure 2 shows the distribution of the trunk cutaneous temperature in a healthy subject. In this case, the

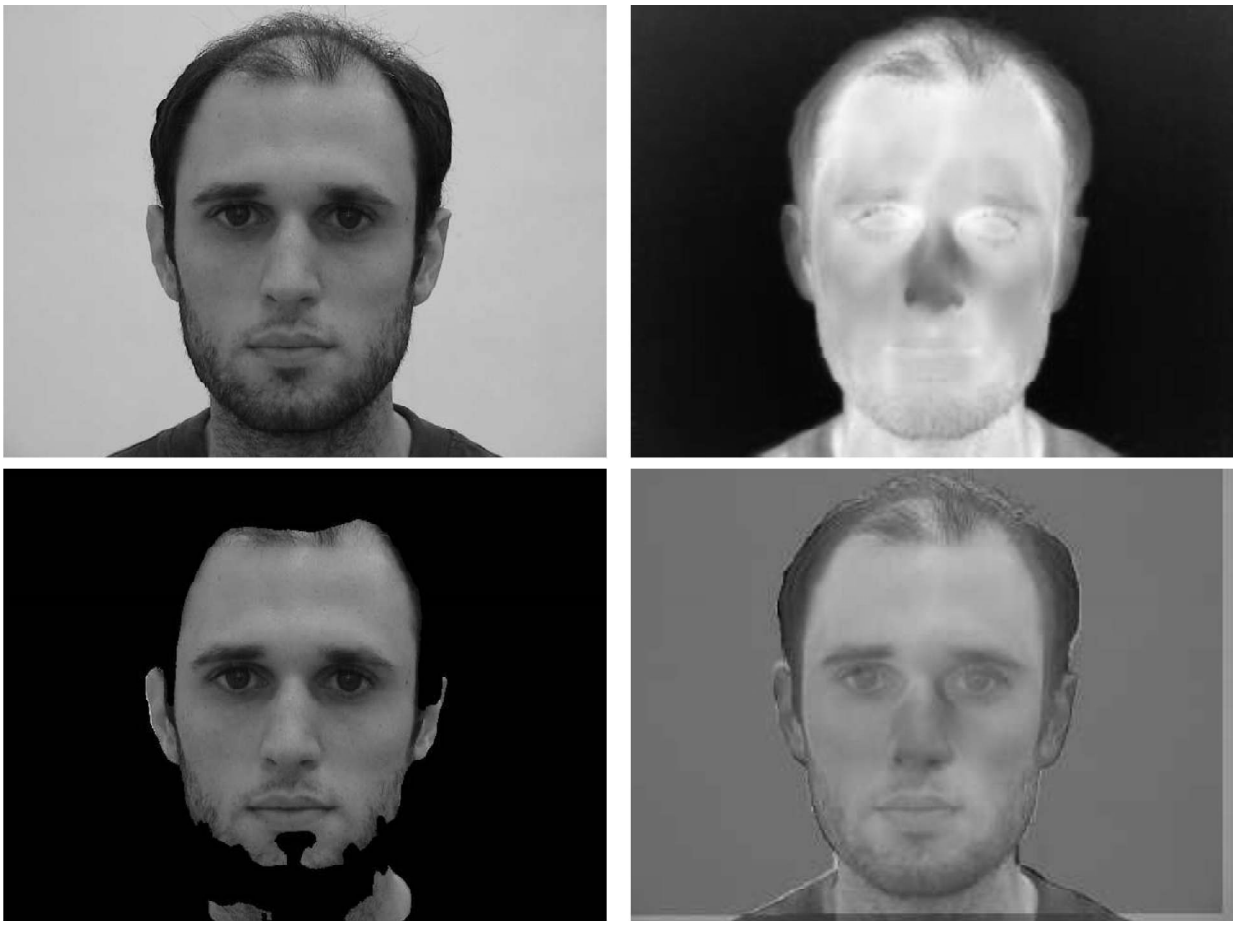

Figure 1: Example 1 of Visual-infrared Overlay: Original Visual Image, Infrared Image, Segmented Visual Image, Composite Image (from Left to Right, Top to Bottom)

Figure 2: Example 2 of Visual-infrared Overlay: Original Visual Image, Infrared Image, Segmented Visual Image, Composite Image (from left to right, top to bottom)

generated images allow to easily associate the temperature distribution of given regions of interest with the subject's anatomy.

Figure 3 shows the distribution of the hand cutaneous temperature in an healthy subject. The generated images permit to localise superficial vessels, which appear either colder or warmer than the surrounding tissue, on the visual image, where they cannot be visually recognised and localised. This capability is extremely helpful in a variety of conditions (fistula or scleroderma, for example) where the clinician needs to access or (non-invasively) monitor the vessel geometry, or in the assessment of morphea (localised scleroderma) patients [6]. Figure 3 also shows how the weights between the two imaging modalities can be set so as to put more or less emphasis on one of the original images.

Finally, Figure 4 gives an example of the application where the user is performing a landmark-rather than an intensity-based registration. Control points are placed on one of the images which are then copied to the other image. Landmarks can be adjusted as a complete set through rotation, translation and scaling operations or individually. Provided the control points are placed correctly, landmark-based registration generates highly accurate overlay images.

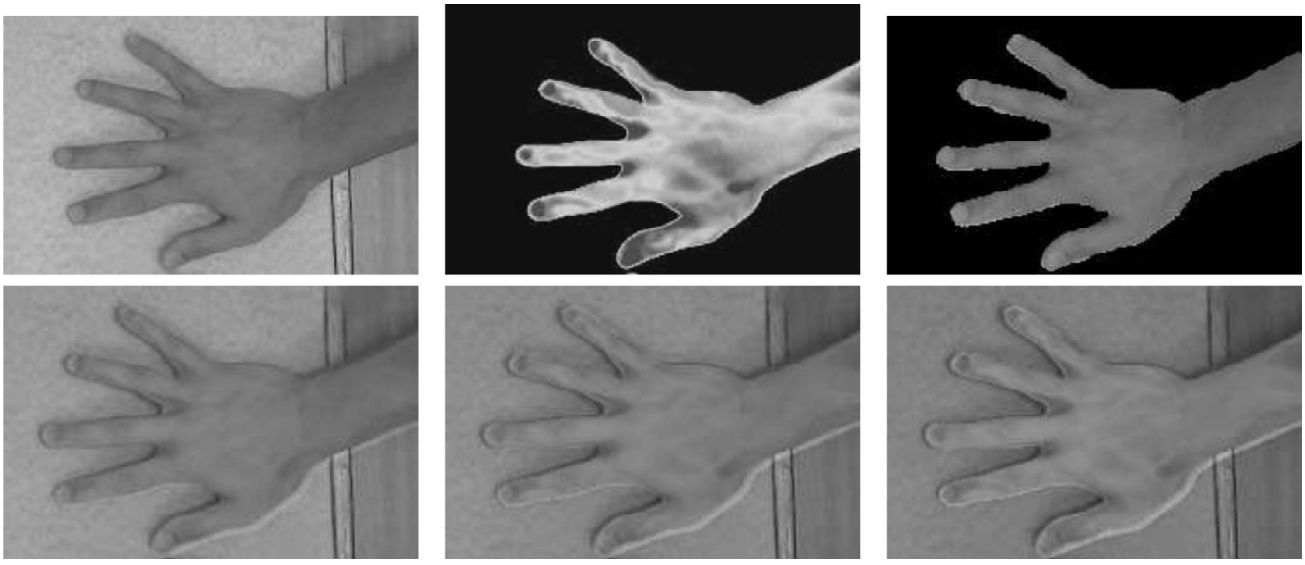

Figure 2: Example 3 of Visual-infrared Overlay. Top row: Original Visual Image, Infrared Image, Segmented Visual Image. Bottom row; three Different Overlays Generated through Different Weightings between Visual and Infrared Images

Figure 4: Examples of Landmark-based Registration. Control Points have been set for the Visual Image and Copied over to the Infared Image

5. CONCLUSIONS

In this paper we have presented our system for superimposing visual and infrared medical images. In a pre-processing step, the image area corresponding to the patient is extracted using a skin detection algorithm. Intensity-based multi-modal registration based on a mutual information is then applied to geometrically align the two images. This process is performed by a set of intelligent agents collaborating via a blackboard architecture. In case the registration fails, the user has the possibility to employ a landmark-based registration method after specification of corresponding control points in both images.

The system is employed at the Institute for Advanced Biomedical Technology at Foundation G. D'Annunzio University in Chieti for monitoring the evolution of the clinical scenario in patients suffering from Raynaud's phenomenon and scleroderma, dermatologic diseases, and recovery from burns. Another important application for which the system is being used is in the field of neuropsychology, for studying emotional responses through autonomic thermal effects and facial expressions.

While at the moment the system is restricted to overlaying static images, in future we will look at incorporating facilities to superimpose dynamic sequences. In addition, we are interested at providing overlays of images that have been taken at different times, typically many months apart, of the same patient. These will for example prove useful in monitoring of how a disease is developing, or whether the current treatment shows any effects.

REFERENCES

[1] N. Anbar, L. Milescu, A. Naumov, C. Brown, T. Button, C. Carly, and K. AlDulaimi. Detection of Cancerous Breasts by Dynamic Area Telethermometry. *IEEE Engineering in Medicine and Biology Magazine*, 20(5): 80–91, 2001.

[2] C. M. Black, K. J. Murray, K. J. Howell, J. Harper, D. Atherton, P. Woo, G. Martini, and F. Zulian. Juvenile-onset Localized Scleroderma Activity Detection by Infrared Thermography. *Rheumatology*, 41(10): 1178–1182, 2002.

[3] P. Ekman and W. Friesen. *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Consulting Psychologists Press, Palo Alto, 1978.

[4] M. M. Fleck, D. A. Forsyth, and C. Bregler. Finding Naked People. In *4th European Conference on Computer Vision*, pages II: 593–602, 1996.

[5] J. F. Head, F. Wang, C. A. Lipari, and R. L. Elliott. The Important Role of Infrared Imaging in Breast Cancer. *IEEE Engineering in Medicine and Biology Magazine*, 19: 52–57, 2000.

[6] K. J. Howell, A. Lavorato, M. Visentin, R. E. Smith, G. Schaefer, C. D. Jones, L.Weibel, C. P. Deton, J. Harper, and P. Woo. Validation of a Protocol for the Assessment of Skin Temperature and Blood Flow in Childhood Localised Scleroderma. *Skin Research and Technology*, 15(3): 346–356, 2009.

[7] B. F. Jones. A Reappraisal of Infrared Thermal Image Analysis for Medicine. *IEEE Trans. Medical Imaging*, 17(6): 1019–1027, 1998.

[8] T. M. Lehmann, C. Gonner, and K. Spitzer. B-spline Interpolation in Medical Image Processing. *IEEE Trans. Medical Imaging*, 20(7): 660–665, 2001.

[9] F. Maes, D. Vandermeulen, and P. Suetens. Comparative Evaluation of Multiresolution Optimization Strategies for Multimodality Image Registration by Maximization of Mutual Information. *Medical Image Analysis*, 3: 373–386, 1999.

[10] J. B. A. Maintz and A. Viergever. A Survey of Medical Image Registration. *Medical Image Analysis*, 2: 1–36, 1998.

[11] A. Merla, L. Di Donato, S. Di Luzio, G. Farina, S. Pisarri,M. Proietti, F. Salsano, and G.L. Romani. Infrared functional imaging applied to Raynaud's phenomenon. *IEEE Engineering in Medicine and Biology Magazine*, 21(6): 73–79, 2002.

[12] A. Merla, L. Di Donato, G. L. Romani, M. Proietti, and F. Salsano. Comparison of Thermal Infrared and Laser Doppler Imaging in the Assessment of Cutaneous Tissue Perfusion in Scleroderma Patients and Healthy Controls. *Int. Journal of Immunopathology and Pharmacology*, 21(3): 679–686, 2008.

[13] A. Merla and G. L. Romani. Functional Infrared Imaging in Medicine: A Quantitative Diagnostic Approach. In *28th IEEE Int. Conference on Engineering in Medicine and Biology*, pages 1: 224–227, 2006.

[14] A. Merla and G. L. Romani. Thermal Signatures of Emotional Arousal: A Functional Infrared Imaging Study. In *29th IEEE Int. Conference on Engineering inMedicine and Biology*, pages 247–249, 2007.

[15] A. Merla, G. L. Romani, S. Di Luzio, L. Di Donato, G. Farina, M. Proietti, S. Pisarri, and S. Salsano. Raynaud's Phenomenon: Infrared Functional Imaging Applied to Diagnosis and Drug Effects. *Int. Journal of Immunopathology and Pharmacology*, 15(1): 41–52, 2002.

[16] N. Otsu. A Threshold Selection Method from Grey-level Histograms. *IEEE Trans. Systems, Man and Cybernetics*, 9(1): 62–66, January 1979.

[17] P. Plassmann and E. F. J. Ring. An Open System for the Acquisition and Evaluation of Medical Thermological Images. *European Journal of Thermology*, 7: 216–220, 1997.

[18] D. Shastri, A. Merla, P. Tsiamyrtzis, and I. Pavlidis. Imaging Facial Signs of Neurophysiological Responses. *IEEE Trans. Biomedical Engineering*, 56(2): 477–484, 2009.

[19] R. Tait, G. Schaefer, A. A. Hopgood, and S. Y. Zhu. Efficient 3-d Medical Image Registration using a Distributed Blackboard System. In *28th IEEE Int. Conference Engineering in Medicine and Biology*, pages 3045–3048, 2006.

[20] R. J. Tait, G. Schaefer, and A. A. Hopgood. Intensity-based Image Registration using Multiple Distributed Agents. *Knowledge-based Systems*, 21(3): 256–264, 2008.

[21] S. Uematsu. Symmetry of Skin Temperature Comparing One Side of the Body to the other. *Thermology*, 1(1): 4–7, 1985.

[22] P. A. Viola and W. M. Wells III. Alignment by Maximization of Mutual Information. *Int. Journal of Computer Vision*, 24(2): 137–154, 1997.

[23] B. Wiecek, S. Zwolenik, A. Jung, and J. Zuber. Advanced Thermal, Visual and Radiological Image Processing for Clinical Diagnostics. In *21st IEEE Int. Conference on Engineering in Medicine and Biology*, 1999.

[24] J. Zitova, B. anf Flusser. Image Registration Methods: A Survey. *Image and Vision Computing*, 21: 977–1000, 2003.