Computer Vision Winter Workshop 2007, Michael Grabner, Helmut Grabner (eds.)
St. Lambrecht, Austria, February 6–8
Graz Technical University

Texture features for affine registration of thermal (FLIR) and visible images

Andreja Jarc, Janez Perš, Peter Rogelj, Matej Perše, and Stanislav Kovačič

University of Ljubljana, Faculty of Electrical Engineering
andreja.jarc@fe.uni-lj.si

Abstract *The aim of our research is to analyse the importance of texture information for affine registration of gray-scale far-infrared (FLIR) images and gray-scale images taken in the visible spectrum. Texture features are extracted by Laws texture coefficients and used for computing registration criterion functions. The proposed feature based approach is compared to the commonly used approach, where a registration criterion function is computed directly from intensity features, i.e. grey values. The experiments on a small database of ten image pairs show that our texture feature based registration method works well for the given image modalities. Furthermore, we can expect more robust and more correct registration when texture based criterion function instead of intensity based one is used.*

1 Introduction and motivation

Thermal imaging is a well known imaging technique. The application domain of thermal imaging has been expanded in the civilian domains, such as surveillance, people tracking and others. For many of those applications, it is beneficial if thermographic images are complemented with visible spectrum images. Sometimes it is critical that images of both modalities are correctly registered – that every pixel pair in visible and thermographic image depicts the same point in the observed scene.

There are several ways to register such images. The precise mechanical adjustment of cameras such that pixels of both cameras would be aligned is difficult. The full and accurate calibration is impractical for surveillance applications, as the dual-modality calibration targets, covering wide area would be needed. Manual calibration depends on the operator's skill to find a number of points which are properly distributed across the observed scene, and well visible in both modalities. The last option, the retrospective registration of images should provide at least some level of automation and is the main topic of this paper.

In the rest of the paper, we will first briefly discuss thermographic image modality and thermographic image formation, along with the reasons why visible to thermographic registration is a challenging problem. Next, we will present some of the related work on multi modality image registration, followed by the description of our methodology and our image dataset. Finally, we will present results, obtained using our methodology on our dataset, and compare them with manual registration, based on points, selected by multiple human operators. The paper is concluded with a discussion and some ideas for the future work.

2 Thermographic image modality

Thermographic images depict the *far infrared* radiation, emitted by the observed scene. It should be noted that far infrared radiation, depicted in thermographic images is quite different from the *near infrared* radiation, which follows the red part of the visible spectrum in the electromagnetic radiation spectrum. In contrast to near infrared images, far infrared images depict not only the reflected radiation but also the radiation emitted from objects. A naive interpretation of thermographic images might be that they depict the temperature of the observed scene. This simple explanation is often inadequate, and for many applications, it is wrong.

2.1 Thermographic image formation

Thermographic images do not depict the temperature of the observed object, they depict the electromagnetic radiation of an object in the far infrared range, which is about $6-15\mu m$. Except for significant difference in wavelength, the far infrared radiation shares many properties of the visible light (which occupies the range of $0.7-0.4\mu m$), since visible light, near infrared and far infrared radiation all behave according to the same physical laws. Consequences of this behavior are interesting and are detailed in [9].

Usually, the object does not emit a radiation of a single wavelength. However, the peak wavelength λ can be obtained by differentiating the Planck's formula, and is given by the Wien's formula [1]:

$$\lambda = \frac{2898}{T}\,[\mu m]\,, \tag{1}$$

where T represents the object temperature on a Kelvin scale. It turns out that the objects at the room temperature (300 K) emit the wavelengths with the peak at $9.7\mu m$, which is in the middle of far infrared range. This is the main motivation for the far infrared imaging: *in this range of electromagnetic radiation, objects at room temperature appear to "glow" without any external source of light.*

Although the thermographic image represents both the emitted and reflected energy the largest contribution comes from the emitted energy. The relation between the temperature of an object and the emitted energy is given by the Stefan-Boltzmann law [1]:

$$W = \varepsilon\sigma T^4\left[Watt/m^2\right], \tag{2}$$

where T again represents temperature, W represents energy, σ represents Stefan-Boltzmann constant, and ε represents the object's spectral emissivity. Emissivity is the property of the material from which is the object made and varies with wavelength. Emissivity is in the range $[0..1]$, and describes the percent of the radiation that the object emits with respect to the ideal blackbody (which has the $\varepsilon = 1$). Thus, since the emissivity varies with the material, it is the important factor in thermographic image formation.

2.2 Problems in thermal-to-visible image registration

It can be seen that the contrast in thermographic images may be the result of either *different temperatures of different objects on the scene* or of *different emissivities of different objects with same temperature* – or both, which is the most frequent case. In the visible spectrum, the relations are entirely different, since the main contribution to the brightness of the image pixels comes from the reflected light.

Since it is obvious that entirely different factors influence the pixel intensities in both image modalities, the pixel intensities in thermographic and visible spectrum images have *in general no direct relationship.*

Moreover, thermographic images are usually of lower resolution and with less pronounced edges. Edges in thermographic images may form due to the transition between different materials, or due to transition between different temperatures. The latter are usually gradual and therefore difficult to use in registration.

Due to all described problems, many well established methods for registration of images work poorly or do not work at all. For example, the publicly available demo [8] for well known SIFT descriptors, proposed by Lowe [7] for registration of visible spectrum images is unable to find any correspondences on most of the thermographic image pairs from our database.

3 Image registration using texture features

Due to radically different way of image formation in visible spectrum images and thermographic images, an intensity mapping between images cannot be used to solve the registration problem in general. In the literature some attempts to register multi-sensor images invariant to intensity dissimilarities are found. The multi-sensor image registration can broadly be classified into two major classes of algorithms. First methods use invariant image representation (i.e. feature based registration). By invariant image representation it is meant a representation that is invariant to changes in brightness and contrast, as well as to contrast reversal. Some examples of invariant image representations are edge maps [2] [3], contour features [5] and point features [6]. However, the process of creating an invariant image representation is usually a tedious and subjective task and, furthermore, difficult to automate.

The second class of methods use a multi-modality similarity measure which allows arbitrary mapping between image intensity values and therefore do not require an invariant image representation. One of such measures is mutual information (MI), which is a measure of information between two images. The method is usually applied directly to raw multi-sensor images [10].

However, MI-based similarity measure considering only intensity values can result in misregistration, due to several issues. Mapping between intensity values depends on several characteristics of an object: emissivity, temperature and color. MI is based on global estimation of the intensity mapping and therefore cannot model those local variations.

In this paper we present an alternative approach to multi-sensor image registration, based on texture features. This approach couples advantages from the methods of invariant image representation and use of multi-modality similarity measure.

The mapping between texture features has a global nature and therefore enables the use of MI as a similarity measure. Moreover, this method can be easily automated.

Texture feature based image registration is rarely seen in the literature, but some experiments show that it delivers promising results in the case of poor quality images [12].

Texture feature based registration requires, first, extraction of texture features from both of the images. In our approach we used Laws texture coefficients. Laws [4] developed a set of two-dimensional filter masks, which are composed of combinations of four one-dimensional filters. Each of the one-dimensional filters extracts a certain texture feature from an image. These features are: level (L), edge (E), spot (S) and ripple (R). The 1-D filter masks are the following:

L5=[1 4 5 4 1]
E5=[-1 -2 0 2 1]
S5=[-1 0 2 0 -1]
R5=[1 -4 6 -4 1].

By convolving these filter masks with each other, a set of symmetric and anti-symmetric filters are obtained. The obtained 5x5 combinations of filter masks are:

L5L5	E5L5	S5L5	R5L5
L5E5	E5E5	S5E5	R5E5
L5S5	E5S5	S5S5	R5S5
L5R5	E5R5	S5R5	R5R5

Laws claims that the most useful texture features are E5L5, S5E5 and S5L5 and their transposed pairs [4]. Our analysis was limited only to the proposed texture features. The same texture features were extracted from both, visible spectrum image and FLIR image, because the registration can only be driven by information, that is common to both images involved in the registration process.

Each filtered image was converted to a texture energy image. Local texture energy is measured by the sum of absolute values in a window, or local region, of the filtered image. We used a 15 x 15 pixel moving window for the texture image transform. The mean value and standard deviation (σ) of texture energy image were calculated and only the values between $\pm 3\sigma$ were retained. Finally, the retained texture energy values were scaled from 0 to 255 integer levels yielding 8-bit quantization.

Our criterion functions were computed by measuring MI of the quantized texture energy image:

$$MI(A,B) = H(A) + H(B) - H(A,B), \quad (3)$$

where $H(A)$ and $H(B)$ are the Shannon entropies of image features for both of the images and $H(A,B)$ is their joint entropy.

Entropy $H(\cdot)$ is computed as:

$$H(\cdot) = -\sum_i p(i) \cdot log_2 p(i), \quad (4)$$

where *p* is a probability distribution of features on an image. As mentioned earlier, MI is also commonly used as a criterion function for multi-modality intensity feature registration. It was first introduced to the field of image registration in 1995 [**?**]. A survey of MI based image registration is found in [10]. From the definition of MI in Eq. 3 high mutual information value corresponds to good image alignment. Thus, the registration procedure must optimize transformation parameters in such a way that MI value is maximized.

After the texture features E5L5, S5E5 and S5L5 were extracted from image pairs, the features which were the most significant were used for registration. The choice of features was performed manually. In the majority of cases the E5L5 texture features were used, but in some also S5E5 and S5L5 textures were used for registration. To simplify the notation, E5L5-L5E5, E5S5-S5E5, S5L5-L5S5 symbolise the texture images obtained from the corresponding texture features, respectively.

The registration procedure where the criterion function was maximized was carried out by Powell's optimisation method [11].

The correctness of texture based criterion function for registration of multi-sensor images was compared with intensity based criterion function. In this case MI was measured from 8-bit image intensity values of overlapping image regions.

4 Experiments

To test the performance of texture based image registration, we performed tests on a small database of ten visible-thermographic image pairs. Thermographic images were captured using E25 thermal camera from FLIR Systems AB, which can be used both to provide live video and to capture images to internal memory. The image sensor of E25 has the resolution of 160×120 image pixels, however, the camera produces JPEG images of double that size. Images were captured with grayscale palette (black = cold, white = hot) with automatic scaling - the camera determined the palette levels for every image separately, based on the minimum and maximum thermographic temperature in each image. These images were resampled to the final resolution of 160×120 image pixes.

Visible spectrum images were captured in color using Olympus 2100 digital camera, converted to grayscale and resampled to the final resolution of 160×120 image pixels. Figures 1 and 2 show the experimental database.

In the first test, all image pairs were registered by use of texture features, and in the second test, registration was performed by use of intensity features.

In order to evaluate the registration results the comparison with the estimate of ground truth was performed. The true ground truth for registration of our image pairs was not known. To estimate the performance of our method we used six estimations of correct image registrations, contributed by six independent observers. Each estimation was obtained by manually selecting on average 7 corresponding points per image.

The registration evaluation experiment was designed as follows:

- Initially, images were not registered. The mean Euclidean distance (i.e. the average Euclidean distance, considering six independent observers) of manually selected points on the reference and floating images was computed, see the results in Table 1, notation: $Unreg$.
- From the first three manually selected points an affine transformation model was computed for each image pair. The rest of the manually selected points were transformed with the obtained affine transformation and the mean Euclidean distance between the points was computed. The error distance tolerance for selecting the points was set to < 2 pixels. If the required tolerance was not achieved, the operator was asked to select points again until the requirement was fulfilled as close as possible, see the results Table 1, notation: $Manual$.
- The manually selected points were transformed with affine transformation model obtained as a result of texture feature based registration. The mean Euclidean distance between transformed manually selected points was estimated, see the results in Table 1, notation: $Ttex$.
- Finally, the manually selected points were transformed with affine transformation model obtained as a result of intensity feature based registration. The mean Euclidean distance between transformed manually selected points was estimated, see the results in Table 1, notation: $Tint$.

5 Results and discussion

The mean Euclidean distances of corresponding points for each image pair are shown in Table 1 in the same order as described above.

Performance of the texture feature based registration depends on image contents - some image pairs can be registered easily, and for some the method does not work well ($img002$ and $img008$). The reason of registration failure are similar local texture representations which increase the risk of local extrema attraction. Furthermore, in the case of $img002$ all of the six manual operators who were selecting the points for manual registration had problems in selecting points with sufficient accuracy. This implies that $img002$ is a borderline case for which registration is barely possible.

From the results it can be concluded that the texture feature based registration works well except for the above mentioned examples. When those two examples are discarded the error distance is 3.29±0.40 pixel on average. For manually registered images the error distance is 1.45±0.36 pixel on average. It can be seen that the error of texture feature

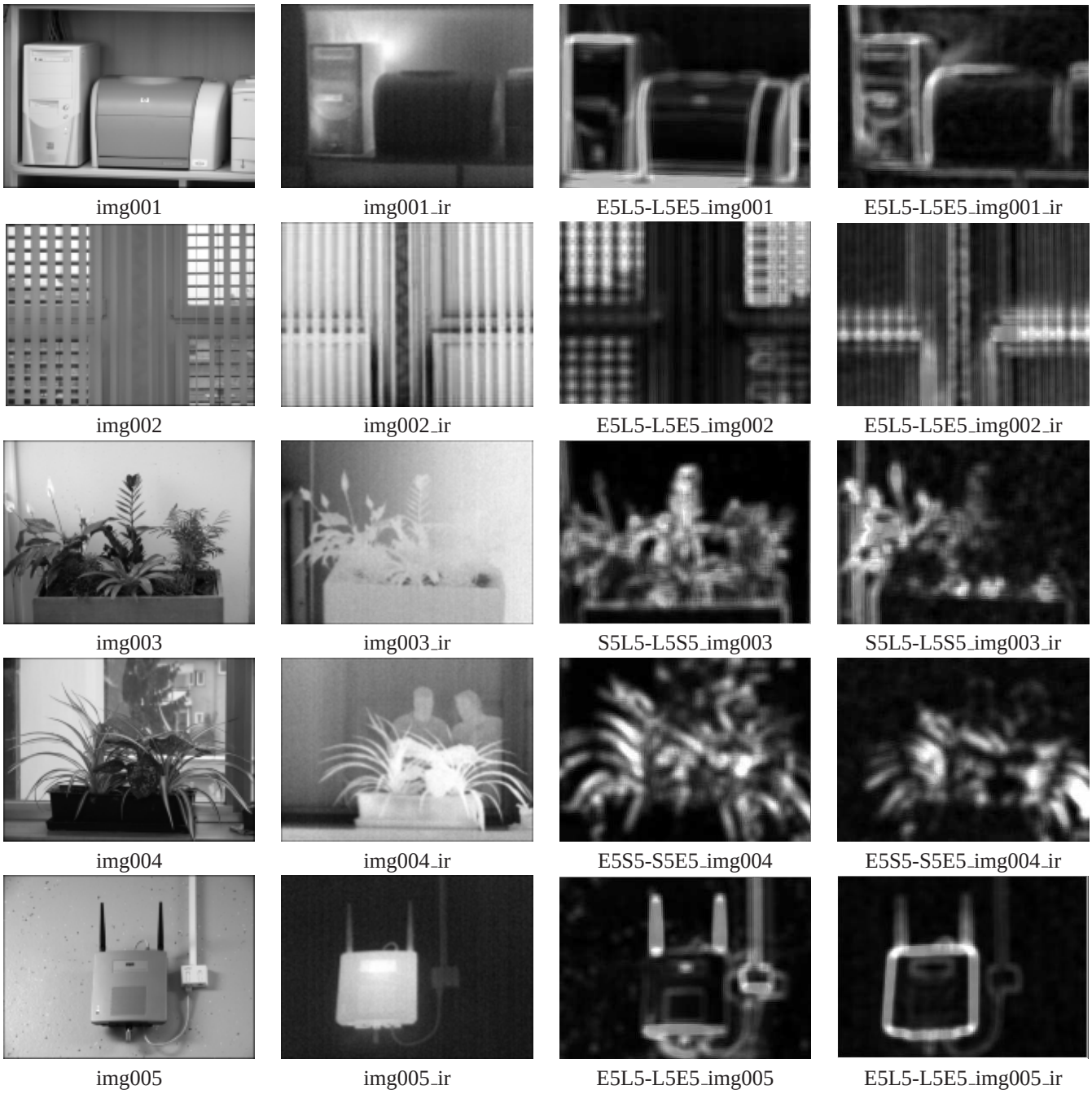


Figure 1: Experimental and texture image pairs 1-5. Visible spectrum images are in the first column (left to right), thermal images are in the second column, texture images derived from visible spectrum images are in the third column and texture images derived from thermal images are in the fourth column.

based registration is about twice the error of the manual registration.

In order to show that the proposed approach is indeed better than intensity based registration, we performed the comparison with intensity based registration. The results are shown in the fourth column in the Table 1. It can be seen that except for $img006$ the intensity based approach delivers significantly higher errors than texture based approach.

The probable explanation for the better performance of the texture feature based approach with respect to intensity based approach is that information which is useful for registration lies only in image structure, while image intensities by themselves may be misleading.

It is interesting that one of the images ($img006$) is surprisingly well registered by the use of intensity features. The reason why the intensity based registration works well is that there are no local variations of relationship between image intensities.

It should be noted that texture feature selection – which is a manual step in our method – significantly influences the registration results. Not all available texture features extract relevant information for registration of a particular image

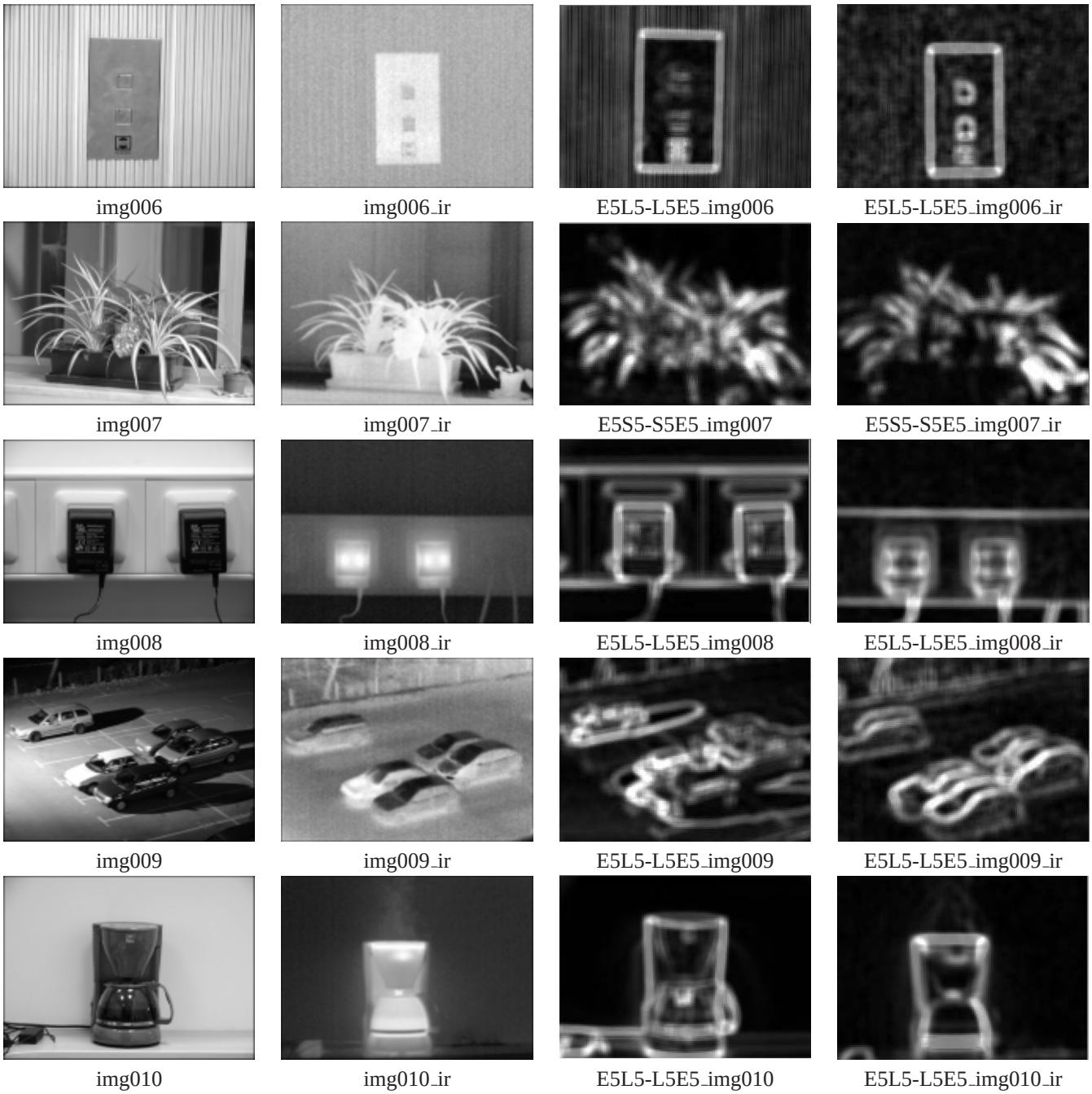


Figure 2: Experimental and texture image pairs 5-10. Visible spectrum images are in the first column (left to right), thermal images are in the second column, texture images derived from visible spectrum images are in the third column and texture images derived from thermal images are in the fourth column.

pair. Therefore, it is critical that the most appropriate features are chosen for each image pair individually.

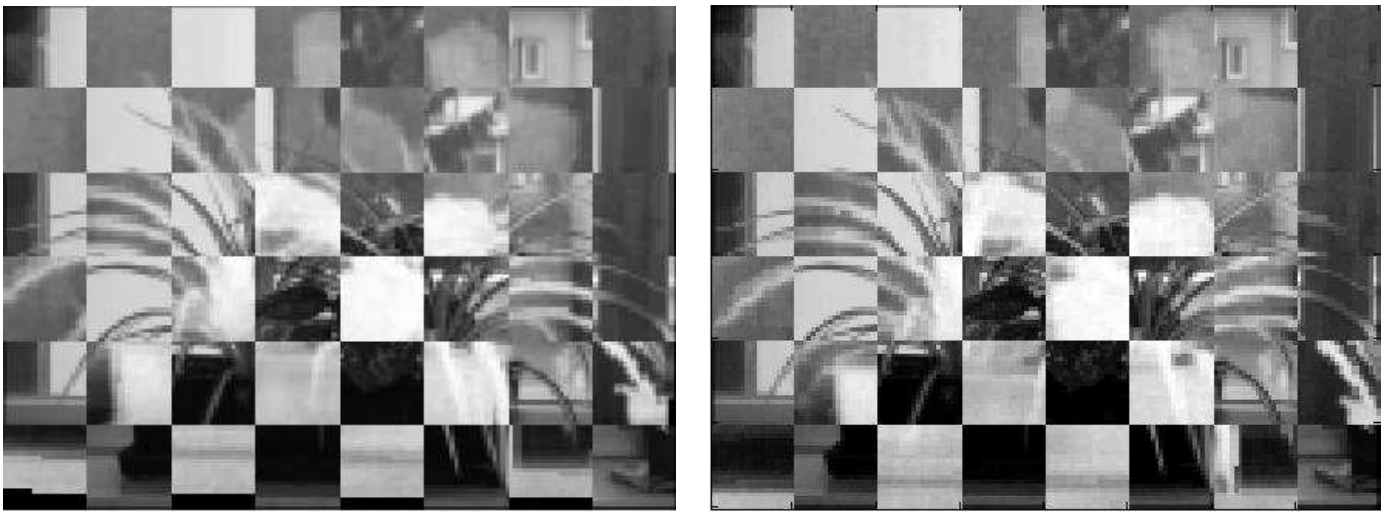
Regarding all of the above, the texture feature based registration method delivers promising results, even though it is not fully automatic. During the experimental work, we found out that even perfect manual registration is difficult as the choice of the right points for manual registration requires a lot of experiences in observing thermographic images. Either manual adjustment of parameters or the manual selection of the corresponding image points are difficult for an operator not trained in dealing with thermal images. Therefore, our registration method can be used for any given pair of thermographic and visible spectrum images, but still the manual inspection should follow the registration.

6 Conclusion

In the present paper we propose a texture-based approach to multi-sensor image registration by extraction of texture features which yield some extra information about local structure representations. It has been shown that by the use of the criterion function that involves texture features, a better matching (in comparison to intensity feature registra-

Table 1: Mean Euclidean distances of the corresponding points from thermographic and visible spectrum images. Results are expressed in pixels as distance ± standard deviation.

Image pair				
$Image$	Unreg.	Manual	Ttex	Tint
img001	6.44 ±0.60	1.29± 0.47	3.29± 0.21	10.79± 0.39
img002	13.18±1.36	1.13± 0.21	12.66± 1.82	15.74± 1.36
img003	9.00 ±0.39	1.75± 0.26	3.56± 0.30	6.30± 0.34
img004	16.17±1.65	2.13± 0.50	2.86± 0.40	7.44± 1.24
img005	18.73±0.28	1.24± 0.20	2.69± 0.35	61.57± 0.69
img006	7.75 ±0.29	0.96± 0.27	4.32± 0.32	0.81± 0.11
img007	3.87 ±0.53	1.49± 0.24	2.25± 0.37	9.63± 0.49
img008	26.12±2.21	1.34± 0.27	11.05± 1.20	17.5± 1.93
img009	7.67 ±0.42	1.18± 0.59	3.13± 0.52	11.12± 0.71
img010	9.61 ±0.86	1.58± 0.33	4.19± 0.75	14.47± 0.70



Texture based registration Gold standard registration

Figure 3: Registered image pair 004. Thermal image is transformed to visual spectrum image coordinates and shown as checkerboard. Texture based image registration in this case delivers results which are similar to gold standard registration.

tion) of the given images may be achieved without presegmentation.

Registration by use of texture features shown in this paper is still not fully automatic as it requires manual feature selection. In the future work we will focus on automatical selection of features optimal for registering a given pair of images. Furthermore, several extracted features could also be combined into new features used in the registration process.

Acknowledgement

This research was supported in part by Slovenian Research Agency (ARRS) contracts P2-0232, 3211-05-000557, Z2-7570, and Slovenian Ministry of Defence (MORS) contract CiVaBis M2-0156.

References

[1] *ThermaCam Reporter 2002: Operator's Manual*. Flir Systems AB, 2002.

[2] Kristin Dana and P Anandan. Registration of visible and infrared images. In *SPIE Conf. on Arch., Hardware and FLIR in Auto. Targ. Rec.*, pages 1–12, 1993.

[3] Yosi Keller and Amir Averbuch. Multisensor image registration via implicit similarity. *IEEE Transactions on pattern analysis and machine intelligence*, pages 794–801, 2006.

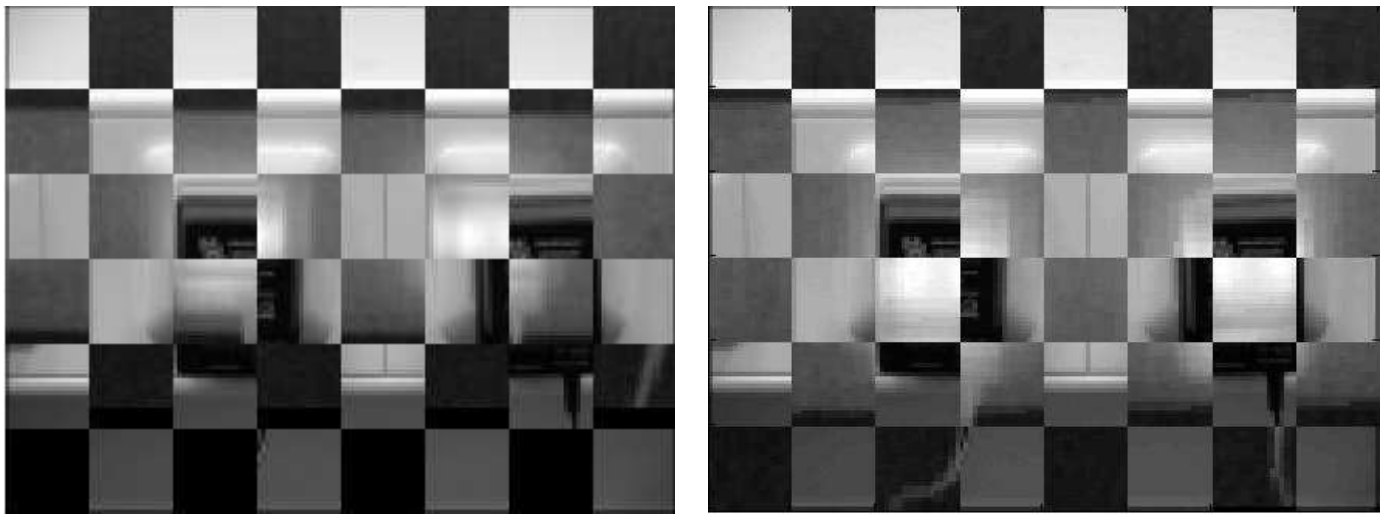
[4] Kenneth Laws. Rapid texture identification. In *In SPIE Image Processing for Missle Guidance*, pages 376–380, 1980.

[5] H Li, BS Manjunath, and SK Mitra. A contour-based approach to multisensor image registration. pages 320–334, 1995.

[6] H Li and YT Zouh. Automatic eo/ir sensor image registration. In *IEEE Int.Conf. on Image Proc., volume B*, pages 320–334, 1995.

[7] David G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.

[8] David G. Lowe. Demo software: Sift keypoint detector, Version 4, July 2005.



Texture based registration Gold standard registration

Figure 4: Registered image pair 008. Thermal image is transformed to visual spectrum image coordinates and shown as checkerboard. Texture based registration delivers a high error in comparison to gold standard registration.

[9] Janez Pers, Matej Perse, Matej Kristan, and Stanislav Kovacic. Observing human motion using far-infrared (flir) camera – some preliminary studies. In *Thirteenth International Electrotechnical and Computer Science Conference ERK 2004, Portoroz Slovenia*, pages 187–190, 2004.

[10] Josien Pluim, Antoine Maintz, and Max A. Viergever. Mutual information based registration of medical images: a survey. *IEEE Transactions on medical imaging*, pages 986–1004, 2003.

[11] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes in C*. second edition, 1996.

[12] F Rousseau, R Fablet, and C Barillot. Density based registration of 3d ultrasound images using texture information. *Electronic Letters on Computer Vision and Image Analysis*, pages 1–7, 2000.