TOWARDS OBJECTIVE QUALITY ASSESSMENT OF IMAGE REGISTRATION RESULTS

Birgit Möller, Rafael Garcia
Department of Electronics, Informatics and Automation
University of Girona, Campus Montilivi, 17003 Girona, Spain
birgit.moeller@informatik.uni-halle.de, rafa@silver.udg.es

Stefan Posch
Institute of Computer Science, Martin-Luther-University Halle-Wittenberg
Von-Seckendorff-Platz 1, 06099 Halle/Saale, Germany
stefan.posch@informatik.uni-halle.de

Keywords: Image Registration, Quality Assessment, Objective Metric, Voting Schemes, Error Classification.

Abstract: Geometric registration of visual images is a fundamental intermediate processing step in a wide variety of computer vision applications that deal with image sequence analysis. 2D motion recovery and mosaicing, 3D scene reconstruction and also motion detection approaches strongly rely on accurate registration results. However, automatically assessing the overall quality of a registration is a challenging task. In particular, optimization criteria used in registration are not necessarily closely linked to the final quality of the result and often show a lack of local sensitivity. In this paper we present a new approach for an objective quality metric in 2D image registration. The proposed method is based on local structure analysis and facilitates voting-techniques for error pooling, leading to an objective measure that correlates well with the visual appearance of registered images. Since observed differences are furthermore classified in more detail according to various underlying error sources, the new measure not only yields a suitable base for objective quality assessment, but also opens perspectives towards an automatic and optimally adjusted correction of errors.

1 INTRODUCTION

Image sequence processing is one of the basic ingredients in modern vision-based problem solving. The analysis of sequences acquired with active, i.e., non-stationary or rotating cameras, is especially important, offering large flexibility for solving a wide variety of different tasks in a straight-forward and elegant way. When analysing sequences, one of the first fundamental steps is usually the identification of image contents that are shared between different images of a sequence, resulting from image overlaps. In particular, mappings between pixels that represent 2D projections of the same 3D scene point in pairs of images are reconstructed. In the literature this identification of corresponding pixels is usually referred to as *2D geometric image registration*.

The registration of a pair of images is based on a geometric transformation to align corresponding pixels. It is usually provided in terms of a parameterised model that describes the camera motion between the images and its effects on the geometric appearance of each image (Hartley and Zisserman, 2004). During registration the parameters are estimated according to a suitable goal function that quantifies the differences between the images and is to be minimised.

Due to the high relevance of 2D registration in present-day computer vision over the years huge numbers of registration techniques have emerged (Zitová and Flusser, 2003). Generally, goal functions are either *featureless*, i.e., based on image intensity residuals or correlation measures, or *feature-based*, i.e., relying on Euclidean distances between the reprojections of selected image features. Sometimes algorithms working in the frequency domain or resting upon a probabilistic fundament are also applied.

In general, minimising a goal function leads to an optimal image alignment according to the chosen optimisation criterion and the underlying data. The results of 2D registration often yield the basis for more elaborate analysis steps. Consequently, apart from the pure registration result, an objective assessment of the final quality is also of significant interest, i.e., the amount of remaining errors in the pixel-wise alignment of both images. However, this is still a big challenge. Although the various optimisation criteria used within image registration may give valuable cues regarding the final accuracy of the pixel-wise image alignment, they often show a lack of local sensitivity.

As an example, consider the well-known *Mean Squared Error (MSE)* widely used as a distance measure in featureless image registration (cf. Eq. 1). As has been demostrated several times in the literature (Wang et al., 2004), image pairs showing the same MSE may appear completely different according to visual differences and registration quality. This renders an objective evaluation impossible. Consequently, registration quality assessment is nowadays often left to the user for visual inspection. Apart from some works aiming at the analysis of theoretical quality bounds of feature-based registration techniques (e.g., (Weng et al., 1989)), only very few approaches towards an automatic quality assessment have been published so far (e.g., (Kim et al., 2000)). Therefore, more sophisticated metrics are highly recommended.

In this paper we present first steps towards overcoming this lack of existing metrics for an objective assessment of registration quality. We propose a new measure that better correlates with the real quality of a registration result as visually experienced by humans. Our metric primarily aims at an automatic assessment of registration results, subsuming the detection of remaining differences between images. This is an important prerequisite to obtain robust and reliable results in subsequent analysis steps. Perspectively, the metric may also guide the development of enhanced algorithms for image registration itself.

The quality of an image registration result is basically defined by errors within the pixel-wise geometric alignment of two images that cause intensity residuals between the images. Consequently, analysing these residuals yields a suitable starting point for the investigation of quality metrics with high local sensitivity. However, in doing so it turns out that intensity differences have to be interpreted carefully. Not all detected residuals are directly related to registration failures. There are additional underlying sources of errors that also have a considerable impact on the final result. Since each of these sources shows its own characteristics and requires an individual compensation strategy, an explicit separation is inevitable.

Our goal to meet these requirements led to the deployment of a detailed *taxonomy* of differences and underlying sources in image registration that guide the development of our new quality metric. Basically, six classes of differences between registered images were identified and further grouped into two main categories: *registration errors*, which are directly related to the geometric registration, i.e., a misalignment of corresponding pixels, and *visual errors*, that subsume remaining intensity differences between correctly aligned pixels and mainly influence the visual appearance of accurately registered images.

Based on the results of these exploratory studies, our new quality metric aims at accounting for these different error categories during quality assessment. Therefore, it extends common image quality measures like those used within the context of assessing loss of quality in image compression into the new field of image registration. The main advantage of the new metric is a high local sensitivity achieved by focusing on pixel-wise quality criteria. In addition, widely used global averaging schemes are replaced with voting-based strategies for error pooling.

The remainder of this paper is organised as follows. In Section 2 principal ideas of registration quality assessment and the deployed taxonomy of image differences in registration are outlined. Subsequently, the state-of-the-art in image quality assessment is discussed (Sec. 3), yielding the base for our developments. The new voting-based metric is the subject of Section 4. Section 5 presents results of our approach while Section 6 contains conclusions and an outlook.

2 IMAGE DIFFERENCES

The aim of our work is to develop objective metrics for assessing the quality of image registration results. Usually, the quality is directly related to pixel-wise differences between images still observable after model parameter estimation and alignment. We start our investigations with a detailed analysis of effects influencing such differences. This leads to the deployment of a *taxonomy* of sources for image differences in registration which is well-suited to guide the investigation of our new quality metric (Sec. 4).

2.1 A Taxonomy of Image Differences

Differences between registered images are not exclusively related to the result of the registration process. Besides, artefacts that are due to changes in the scene or technical properties of the image acquisition device also appear. These have to be distinguished from effects of the registration. Consequently, in objective quality assessment various classes of sources for differences have to be considered separately. During our experiments six basic classes were identified:

General Registration Failure. General failure of the parameter estimation stage (i.e., due to local minima of the goal function) inducing structural image differences and perspective image misalignment.

Model Failure and Parallax. Motion model selection is usually based on assumptions about camera motion and scene structure; discrepancies from these assumptions may cause (local) misalignment. Regarding 2D motion recovery, *parallax* is an example of such effects resulting from apparent motion of static objects due to camera viewpoint changes.

Lens Distortions. Geometric image distortions resulting from non-linear lens properties that hamper a correct reconstruction of the camera motion.

Vignetting. Non-uniform distribution of energy in an image causing the corners to appear significantly darker than central regions; vignetting mainly results from physical lens and/or CCD properties.

Illumination Changes. Global and local changes of lighting conditions within a scene or in the parameters of the camera that cause corresponding image parts to appear visually different after registration.

Moving Objects. Within non-static scenes the motion model for the camera usually fails to compensate for camera and local object motion at the same time, leading to local inconsistencies between the images.

Each of the classes directly influences the visual appearance of registered images. Related image differences are pairwise independent of each other. Nevertheless, due to the individual properties of each class, two main categories of differences can be distinguished. They have a fundamental impact on the development of metrics for quality assessment. In the following we will refer to the two categories as *registration* and *visual errors*, characterised as follows:

A) **Registration Errors:** This category includes image differences due to erroneous alignments of corresponding pixels. They result from *General Registration Failures*, *Model Failure and Parallax*, or *Lens Distortions* and usually correlate with differences in the local structure of both images.

B) **Visual Errors:** In this case the pixel-wise geometric image alignment is correct, but intensity differences still appear. *Vignetting*, *Illumination Changes* and *Moving Objects* within a sequence are the main reasons for these effects.

Comparing the individual characteristics of both categories it turns out that only errors of the first category are directly related to the registration process. Visual errors are only loosely linked to it. They likewise affect the overall appearance of registered images. However, corresponding artefacts cannot be compensated by improving the registration result, but require an explicit and special treatment.

The distinction between registration and visual errors gives valuable cues regarding the properties that a new metric in registration quality assessment requires. In addition, a more detailed analysis of the various error sources within the two categories supplements important information. In particular, the different members in both categories show significantly varying impacts on the final visual appearance of images. Among registration errors, global misalignment and lens distortions usually affect the whole image, while effects of parallax or minor local misalignments are often restricted to local areas of the images, e.g., to corner regions. Consequently, more attention has to be paid to the detection and correction of lens distortions and global misalignment than to local misalignments. Likewise, among visual errors, global illumination changes and vignetting show a deeper impact on the visual appearance of images than artefacts resulting from moving objects. Figure 1 resumes our resulting taxonomy of differences in image registration. It considers their impact and the necessity for compensation with regard to registration quality and overall visual appearance.

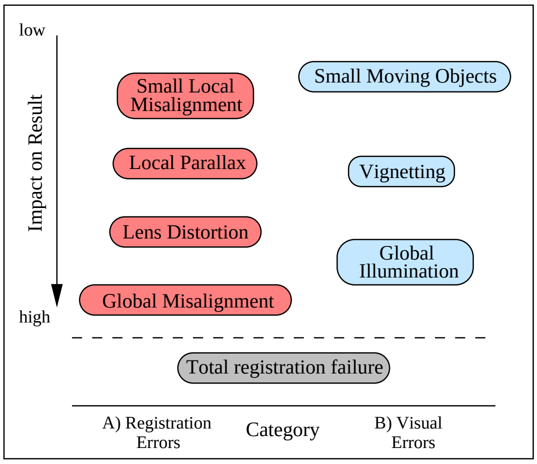

Figure 1: A taxonomy of differences in image registration.

The taxonomy will guide the development of our new metric, discussed in detail in Section 4. Prior to this we will briefly review *image* quality measures used for the general detection of differences between two images, e.g., regarding artefacts resulting from image compression, rendering or moving objects. As the objective assessment of registration quality is deeply linked to these tasks they yield a suitable starting point for our investigations.

3 REVIEW: IMAGE QUALITY ASSESSMENT

Within some research fields, e.g., dealing with the transmission of visual data or photo-realistic rendering (Wang et al., 2004), various quality measures were investigated in the past to quantify differences between two images that show the same contents. One of the first and still widely used measures for quantifying the difference between all pixels u within the overlapping area $\mathfrak{D}_{12}$ of two images I_1 and I_2 is the well-known *Mean Squared Error (MSE)*:

$$\varepsilon_{MSE}(I_1,I_2) = \frac{1}{|\mathfrak{D}_{12}|}\sum_{u\in\mathfrak{D}_{12}}(I_1(u)-I_2(u))^2 \qquad (1)$$

However, since the MSE has shown a lack of discriminative power for various use cases, alternative metrics have emerged. Basically, two principal directions can be distinguished. On the one hand, perceptually motivated measures were investigated in order to emulate the capabilities of the human visual system (HVS). These measures take into account characteristics of the human eye, like masking effects or spectral sensitivity. In addition, environmental parameters influencing the perception of an image are considered, e.g., lighting conditions or viewing angles. Two prominent members of this class are the *Sarnoff Visual Discrimination Model* (Lubin, 1995) and the *Visual Difference Predictor* (Daly, 1993).

On the other hand, quality measures have been developed that put higher emphasis on structural image properties. They usually rely on edge data, local gradient information or image entropy (Xydeas and Petrović, 2000; Wang and Bovik, 2002).

HVS-based measures basically exploit perceptually motivated image properties that influence the visual appearance of an image. On the contrary, the key idea of structural image quality measures is given by the observation that the human visual system focuses strongly on the structural contents of a scene. Consequently, changes within the structural information of an image are assumed to provide a good approximation to perceived image distortions (Wang et al., 2004). Since these data also have an important and stronger impact on the quality of image registration results than perceptually motivated differences, we consider structural metrics as a reasonable basis for our approach.

3.1 Structural Metrics

Most structural image quality measures take a wide variety of structural image properties into account, ranging from mutual information (Qu et al., 2002) to local gradient orientation and magnitude (Xydeas and Petrović, 2000). The recently proposed *Universal Image Quality Index (UIQI)* (Wang and Bovik, 2002) and its generalisation, a *Measure of Structural Similarity (SSIM)* (Wang et al., 2004), basically exploit statistical image properties in terms of average intensity values and correlations related to contrast, luminance and image structure, respectively.

Structural measures are usually calculated pixelwise between the two images I_1 and I_2 to be compared. Depending on the selected structural features, a neighbourhood of each pixel is taken into account, as is done by the UIQI which is exemplarily discussed below. It is defined as follows:

$$\varepsilon^0_{UIQI}(I_1,I_2|w) = \frac{\sigma_{12}}{\sigma_1\sigma_2}\cdot\frac{2\overline{I_1}\,\overline{I_2}}{\overline{I_1}^2+\overline{I_2}^2}\cdot\frac{2\sigma_1\sigma_2}{\sigma_1^2+\sigma_2^2} \qquad (2)$$

ε^0_{UIQI} is calculated within a sliding window w of a fixed size which is shifted over the images. $\overline{I_1}$ and $\overline{I_2}$ denote the average intensity values within each w and (σ_1,σ_2,σ_{12}) refer to correlation coefficients. The pixel-wise calculation of the quality measure results in a map that specifies the local similarity of both images for each pixel. In the context of image data transmission and rendering, these maps are usually used for direct visual evaluation. Alternatively, their entries are summarised by a single numerical value computed from averaging:

$$\varepsilon_{UIQI}(I_1,I_2) = \frac{1}{|W|}\sum_{w\in W}\varepsilon^0_{UIQI}(I_1,I_2|w) \qquad (3)$$

with W being the set of all sliding windows. Structural metrics have proven suitable for quantifying image differences, e.g., with regard to compression techniques and different kinds of image noise (Cadik and Slavik, 2004; Wang et al., 2004). However, regarding the detection and assessment of registration errors, even extensions of the UIQI that put higher emphasis on local image properties during error pooling (e.g., (Cvejic et al., 2005)) show a lack of pixelwise local sensitivity. This is mainly due to the fact that compression artefacts and noise are equally distributed over entire image regions most of the time and require a more region-based evaluation. On the contrary, differences in image registration often show local variations and characteristic spatial patterns that have to be considered explicitly.

Accordingly, new metrics and error pooling strategies with higher local sensitivity are required for registration quality assessment. Within our approach we account for this by carrying out a more detailed local analysis of quality maps that result from the calculation of different structural quality criteria. These appear well-suited for our requirements since high registration quality usually coincides with the preservation of structural properties of both images during the registration process. In the next section we present our approach for a detailed difference classification and an objective quality assessment.

4 VOTING-BASED EVALUATION OF REGISTRATION QUALITY

Our new metric for registration quality assessment focuses on two principal goals. On the one hand we aim at providing a high local sensitivity by putting strong emphasis on pixel-wise calculations, and on the other hand global error pooling strategies are proposed that better preserve pixel-wise information compared to existing structural quality measures. To achieve these goals two main investigations are carried out. Local sensitivity is granted by omitting region-based calculations during the first phase of an evaluation (Subsection 4.1). This strategy yields quality maps with high local sensitivity. Secondly, an improved analysis of the maps is performed. While weighted average values are usually calculated to get a final quality index within state-of-the-art techniques, we propose a voting-based scheme for interpreting the contents of these maps within a global context (Subsection 4.2).

4.1 Local Quality Criteria

We follow the main ideas of structural image quality measures (Subsec. 3.1) and settle our new registration quality metric on structural cues (cf. also (Wang et al., 2004)). Since the overall visual appearance of registered images is at the same time influenced by illumination changes, vignetting or moving objects, these are also taken into account during evaluation.

Given two registered and aligned images I_1 and I_2, pixel-wise features are first calculated according to structural criteria. Three different maps of local measures result from these calculations:

Absolute Intensity Difference Map ***D*****.** Differences within the pixel-wise intensity values of two images always provide cues for possible misalignment. Indeed, the differences themselves are only of minor interest, since they are very sensitive to noise. Hence, we will interpret them in conjunction with other measures to obtain meaningful results (cf. also (Farin and de With, 2005)). The *difference map D* for both images at each pixel position (x,y) within the overlap area of both images is defined as follows:

$$D(x,y) = |I_1(x,y) - I_2(x,y)| \tag{4}$$

Structural Risk Map ***R*****.** One fundamental criterion for assessing the image structure at a certain position (x,y) within an image is the magnitude of the local gradient. We use this value to assess the extent to which single pixels may give reliable cues for registration quality evaluation. At positions where the gradient magnitude is quite small in both images, structure is only weakly distinctive and its analysis may lead to wrong conclusions. Hence, these positions are excluded from structure analysis. With regard to visual errors, however, these pixels give valuable hints, e.g., for identifying vignetting effects, as will be explained in more detail in the next subsection. The binary *risk map R* for marking non-relevant pixels in structure analysis is calculated as follows:

$$R(x,y) = \begin{cases} 1, & \text{if } G_1(x,y) \leq \theta_G \wedge G_2(x,y) \leq \theta_G \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

with G_1 and G_2 being the local gradient magnitudes in both images. θ_G is a suitable threshold. Morphological dilation with a 3×3 squared mask is applied to the risk map to also exclude pixels close to a homogeneous neighbourhood.

Edge Preservation Map ***E*****.** The gradient magnitude is only one component of the structural information given by local derivatives. Also the orientation of the gradient yields valuable cues. In (Xydeas and Petrović, 2000) a metric for image fusion performance was proposed that exploits gradient magnitude and orientation for a perceptually motivated assessment of how well edge information is preserved during image fusion. In our experiments, especially the analysis of gradient orientation has turned out to provide important information for quality assessment. Thus, we use the *edge preservation maps E* proposed in (Xydeas and Petrović, 2000), but solely exploit the gradient orientation of two registered images[1]:

$$E(x,y) = \frac{\Gamma_\alpha}{1 + e^{k_\alpha (A(x,y) - \sigma_\alpha)}}$$

with $A(x,y) = 1 - \frac{|\alpha_1(x,y) - \alpha_2(x,y)|}{\pi/2}$ being defined as the relative local orientation, $\alpha_k = \tan^{-1}\left(\frac{s_k^y(x,y)}{s_k^x(x,y)}\right)$, and $s_k^y(x,y)$ and $s_k^x(x,y)$ being the outputs of convolving image I_k at position (x,y) with horizontal and vertical Sobel templates. The constants in the formulas have been chosen according to the default values suggested by Xydeas et al. with $\Gamma_\alpha = 0.9879$, $k_\alpha = -22$ and $\sigma_\alpha = 0.8$, leading to values of E within the range of $[0,1]$ where 1 indicates absolute similarity.

[1] In the original paper, in addition to both input images the final fusion result is also taken into account.

4.2 Global Assessment

Given the three different maps of local criteria, the overall goal is to automatically assess the global registration quality, and to identify other meaningful image differences. As weighted averaging techniques have not proven to be sensitive enough for our purposes, we propose a *block-based voting strategy* instead. For this the images are divided regularly into 8×8 pixel-sized blocks and for each block the occurrence of registration and visual errors is pixel-wise checked. Given the results of these checks, statistics are computed for each block with the result being a final vote regarding existence and relevance of registration and/or visual errors in the underlying image section. These voting strategies are outlined below.

Registration Error Analysis. According to Section 2, registration errors are due to a misalignment of corresponding pixels between both images and correlate with significant structural differences. We assume a single pixel position (x,y) within registered images to show such a difference if its edge preservation value $E(x,y)$ is smaller than a given threshold θ_E. It is chosen so as to indicate a significant mismatch in the local gradient orientation between both images. Subsequently, a whole block votes for registration errors in the global assessment step if more than 10% of its pixels are fulfilling the threshold criterion.

Visual Error Analysis. In contrast to registration errors, visual errors are mainly related to illumination changes and vignetting (given a correct local geometric alignment of the images). We consider a pixel to show this type of error if the intensity difference between both images exceeds a certain threshold θ_D, provided that the pixel is allowed to vote for visual errors according to the risk map R (Eq. 5), i.e., lying in a homogeneous region of the image. Subsequently, a whole block votes for visual errors if more than 25% of the pixels agree.

Up to now, for each of the blocks within an image, two binary decisions have been taken with regard to the existence of registration and visual errors within the corresponding image section. Given an image where each block is represented by a single pixel, and marking the pixels according to the votes of the blocks, a quality map is obtained. Visual inspection clearly indicates regions where image differences belonging to one of the two categories appear (cf. Fig. 4). However, since we aim at a fully automatic quality assessment that also differentiates between the various sources of errors in both categories, the maps have to be further analysed and interpreted. As widely used (weighted) averaging techniques have shown a lack of local sensitivity we favour a voting strategy for this. Each block participates in two independent voting schemes that subsequently form the foundation for a rule-based decision process to identify the various sources of image differences according to our taxonomy, as explained below.

Final Voting & Decision Rules. The final voting and global quality assessment is done by first counting the number of blocks voting for errors of one or both categories. During the voting the spatial position of the blocks within the image is taken into account since some difference classes are mainly characterised by their spatial distributions (Subsec. 2.1). The blocks are separated into border and central blocks, in which the border of an image is defined to cover approximately the outer third of the image area. In addition, the overall structural information within each block B is considered and characterised by the entropy $H(B)$. H is calculated based on the intensity values v of all pixels (x,y) within a single block:

$$H(B) = \sum_{v=0\ldots255} \frac{1}{p_v} \log p_v \;\; , \;\; p_v = \frac{1}{|B|} \sum_{(x,y)\in B} \delta_{I(x,y),v}$$

A block is considered to contain sufficient structure to vote for registration errors if its entropy normalised to a range of $[0,1]$ exceeds a threshold θ_H.

As a result of these voting processes, we get the ratios of how many central and border blocks are voting for visual and registration errors. They yield the base for identifying the different classes of errors in both categories whereas the following rules are applied:

- *Vignetting* is present between the images if the ratio of border blocks voting for visual errors significantly exceeds the ratio of central blocks.
- *Global Illumination Changes* occur if the ratio of border and central blocks voting for visual errors is more or less the same and exceeds a certain percentage of all blocks voting.
- *Radial Geometric Distortions* force a significantly higher ratio of border blocks to vote for registration errors than that of central blocks.
- *Global Misalignment* results in more or less equally high ratios of border and central blocks voting for registration errors. Contrarily, if there are only very few blocks in total voting for this kind of error, the overall quality seems to be high.

In addition to these rules, any block containing a huge amount of pixels that exceed an intensity difference which is significantly larger than θ_D (compared to neighbouring blocks) may give cues for localising errors due to *Moving Object*, *Local Misalignment* or *Parallax* due to unfulfilled model assumptions.

5 RESULTS

Our approach has been tested on various pairs of registered images. In the following we discuss two representative examples that outline its potential. We focus on effects of lens distortion known to have a significant impact on registration results (Hsu and Sawhney, 1998). Since correcting distortions as a precaution is quite unstable if no assumptions about their impact can be made, detecting them during registration helps to improve the robustness and reliability of the results.

The reference images of both image pairs discussed here are depicted in Figure 2. The image on the left shows a stony yard with a very shallow 3D relief; the right one depicts a map placed on a plane table. Compared to the distance of the camera moving in parallel to the ground, both scenes may be assumed to be planar. In both examples homographies were applied to describe the camera motion which was recovered applying (Mann and Picard, 1996), given about 90% image overlap. The results suffered only from very small misalignments, rendering it difficult to present them adequately here. Thus, we focus on the outcomes of our new metric.

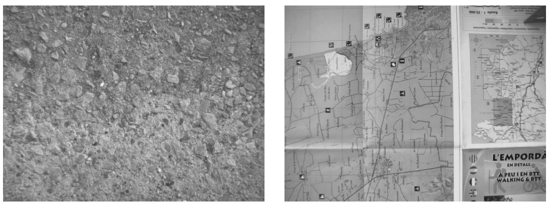

Figure 2: The reference images (640×480 pixels) of the two pairs of registered images exemplary discussed here.

For all evaluations the parameters within our approach were selected as follows: $\theta_G = 5$, $\theta_E = 0.85$, $\theta_D = 2$ and $\theta_H = 0.5$. The resulting registration error map for the first image pair is shown in Figure 3 on the left. Central blocks voting for errors are marked gray, border blocks white, and non-voting blocks as well as blocks without errors are shown in black. It is clearly visible that the blocks voting for registration errors cover the whole image area. Specifically, 62% of the voting border blocks and 58% of the central blocks indicate registration errors. According to the rules defined in Section 4.2 for interpreting the results, this clearly underlines severe problems within the registration. In fact, the result is notably improved after correcting both images for radial lens distortions prior to the registration (Fig. 3, right). We utilised the *Camera Calibration Toolbox*[2] for Matlab where the resulting distortion coefficients are indeed significant: $kc_1 = -0.19$, $kc_2 = 0.21$ and $kc_5 = 0.0001$. After correction only 11% of the border blocks and 6% of the central blocks still voted for registration errors. It should be noted that although lens distortions turned out to be the main reason for the initial registration failure the errors were not exclusively concentrated along image borders. Moreover, a radial symmetric pattern of blocks with errors was observed (Fig. 3).

[2] http://www.vision.caltech.edu/bouguetj/calib_doc/.

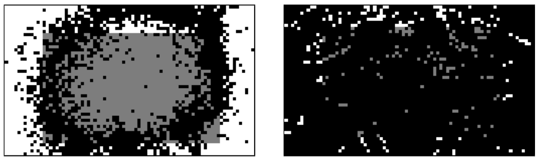

Figure 3: Maps of blocks voting for registration errors in the first example, before (left) and after (right) compensation for lens distortions. Central blocks are marked gray, border blocks white, and non-voting and error-free blocks black.

The results of the quality assessment for the second pair of images are depicted in Fig. 4. The voting results for registration (top) and visual (bottom) errors, given the original images, are shown on the left, and the results after distortion correction on the right.

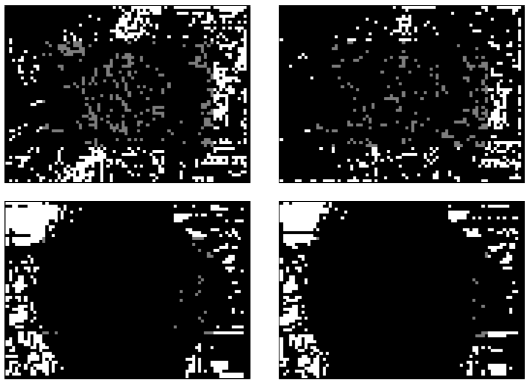

Figure 4: Result maps for the second example, again before (left) and after (right) lens distortion compensation. The maps for registration errors are depicted on top while the results for visual errors are given at the bottom.

Specifically, before lens distortion compensation 50% of the border and 32% of the central blocks voted for registration errors, while only 5% of the central blocks voted for visual errors compared to 70% along the borders. The high ratio of blocks voting for registration errors indicates an overall low registration quality. In this case the slightly higher percentage in outer image parts may indicate problems due to lens distortion. Indeed, after correction, the ratios in the outer and central parts decreased to 24% and 21%, respectively. However, since the ratios are still significant compared to the final ones of the first example, additional factors causing problems in registration and leading to a low quality are probabilistic, e.g., the more pronounced 3D relief of the scene.

With regard to visual errors there is a significant influence of vignetting which is also clearly visible from the original images (Fig. 2, right). Note that the image differences resulting from these effects remain unchanged while compensating for lens distortions as indicated by the final ratios of 65% and 4% after the second registration run. This is as expected since moderate vignetting usually does not affect the registration process, but merely the final visual appearance of the images. Consequently, individual correction steps are required, but not carried out in this case.

6 CONCLUSION

An objective assessment of 2D image registration quality is a challenging task. As common measures for image quality have proven not to be suitable for the special requirements of errors in image alignment, a new metric for this is proposed. Promising results are obtained by exploiting local structural properties of registered images and preserving this information in error pooling by applying voting-based strategies. The indicated registration quality correlates well with the visual appearance of the images and various classes of differences can be distinguished. This capability is of significant importance with regard to subsequent processing steps that aim at an automatic improvement of the results, since different error sources require individual compensation strategies.

While the obtained results outline the high potential of this approach, perspectives for further refinements were also discovered. Sometimes the distribution of blocks voting for registration errors do not clearly indicate the underlying error sources. We plan to tackle this problem by refining the spatial classification of the blocks and by taking global patterns into account. In addition, with regard to local differences resulting from moving objects or parallax detailed examinations of extraordinary high intensity differences will be carried out. Finally, so far the approach relies on various manually adjusted thresholds. Presumably these can be chosen appropriately according to actual image contents, leading to a fully automatic and flexible approach for registration quality assessment.

ACKNOWLEDGEMENTS

This work was supported by a fellowship within the Postdoc-Programme of the German Academic Exchange Service (DAAD), and has also been partially funded through the MOMARNET EU Research and Training Network project (MRTN-CT-2004-505026), and by the Spanish Ministry of Education and Science under grant CTM2004-04205.

REFERENCES

Cadik, M. and Slavik, P. (2004). Evaluation of two principal approaches to objective image quality assessment. In *Proceedings of the 8th International Conference on Information Visualisation (IV)*, pages 513 – 518.

Cvejic, N., Loza, A., Bull, D., and Canagarajah, N. (2005). A novel metric for performance evaluation of image fusion algorithms. *Transactions on Engineering, Computing and Technology*, 7:80–85.

Daly, S. (1993). The visible difference predictor: An algorithm for the assessment of image fidelity. *Digital Images and Human Vision*, pages 179–206.

Farin, D. and de With, P. (2005). Misregistration errors in change detection algorithms and how to avoid them. In *Proc. of IEEE International Conference on Image Processing*, pages II:438–441.

Hartley, R. and Zisserman, A. (2004). *Multiple View Geometry in Computer Vision*. Cambridge University Press.

Hsu, S. C. and Sawhney, H. S. (1998). Influence of global constraints and lens distortion on pose and appearance recovery from a purely rotating camera. In *4th IEEE Workshop on Applications of Computer Vision (WACV)*, page 154 ff., Princeton, NJ, USA.

Kim, H.-S., Kim, H.-C., Lee, W.-K., and Kim, C.-H. (2000). Stitching reliability for estimating camera focal length in panoramic image mosaicing. In *Int. Conf. on Pattern Recognition*, pages 1:596–599, Barcelona, Spain.

Lubin, J. (1995). A visual discrimination model for imaging system design and evaluation. In Peli, E., editor, *Visual Models for Target Detection and Recognition*, pages 245–283. World Scientific, Singapore.

Mann, S. and Picard, R. (1996). Video orbits of the projective group: A new perspective on image mosaicing. Technical Report 338, MIT Media Laboratory Perceptual Computing Section, Boston, USA.

Qu, G., Zhang, D., and Yan, P. (2002). Information measure for performance of image fusion. *Electronics Letters*, 38(7):313–315.

Wang, Z., Bovik, A., Sheikh, H., and Simoncelli, E. (2004). Image quality assessment: From error visibility to structural similarity. *IEEE Trans. on Image Processing*, 13(4):600–612.

Wang, Z. and Bovik, A. C. (2002). A universal image quality index. *IEEE Signal Proc. Letters*, 9(3):81–84.

Weng, J., Huang, S., and Ahuja, N. (1989). Motion and structure from two perspective views: Algorithms, error analysis, and error estimation. *IEEE Trans. on Patt. Anal. and Mach. Intel.*, 11(5):451–476.

Xydeas, C. and Petrović, V. (2000). Objective image fusion performance measure. *Electr. Letters*, 36(4):308–309.

Zitová, B. and Flusser, J. (2003). Image registration methods: a survey. *Image and Vision Computing*, 21(11):977–1000.