Vision, Modeling, and Visualization (2010), pp. 1–8
Reinhard Koch, Andreas Kolb, Christof Rezk-Salama (Eds.)

# Direct Image Registration without Region of Interest

F. Brunet[1,2] and A. Bartoli[1] and N. Navab[2] and R. Malgouyres[3]

[1]ISIT, Université d'Auvergne, Clermont-Ferrand, France
[2]CAMPAR, TU München, Germany
[3]LIMOS, UMR 6158, Clermont-Ferrand, France

**Abstract**

*Standard direct image registration consists in estimating the geometric warp between a source and a target images by maximizing the photometric similarity for the pixels of a Region of Interest (*RoI*). The* RoI *must be included in the real overlap between the images otherwise standard registration algorithms fail. Determining a proper* RoI *is a hard 'chicken-and-egg' problem since the overlap is only known after a successful registration. Almost all algorithms in the literature consider that the* RoI *is given. This is generally either inconvenient or unreliable.*
*In this paper we propose a new method that registers two images without using a* RoI*. The key idea of our method is to consider the off-target pixels as outliers. We define the off-target pixels as those pixels of the source image mapped outside the target image by the current warp. We use the classical robust M-estimation framework to handle both the off-target pixels and the usual outliers caused, for instance, by occlusions. With our formulation, the true image overlap is defined as the set of inliers.*
*Experiments on synthetic and real data with the homography and Free-Form Deformation show that our method outperforms standard approaches in terms of accuracy and robustness while precisely retrieving the overlap in the source and target images.*

Categories and Subject Descriptors (according to ACM CCS): I.4.3 [Computer Graphics]: Image Processing and Computer Vision—Registration

## 1. Introduction

Image registration is the problem of determining a warp that aligns a source image and a target image. It is a problem of wide interest in computer vision and medical imaging, with applications such as image mosaicing [IA99], super-resolution [HBA97], and tracking [HPN99]. The warp is for instance a homography or a Free-Form Deformation.

There are two main approaches to image registration [Sze06]: the feature-based and the direct (or pixel-based) approaches. The feature-based approach [TZ99] estimates the warp parameters from feature correspondences. The direct approach [Bar08, IA99], that we follow in this paper, is to minimize a colour discrepancy measure. For standard direct algorithms, this measure is computed over a given set of pixels $\mathfrak{R}$ called the Region of Interest (RoI) [BGM04, Bar08]. Let $\mathcal{W} : \mathbb{R}^2 \times \mathbb{R}^p \rightarrow \mathbb{R}^2$ be a warp parametrized by a vector $\mathbf{p} \in \mathbb{R}^p$. In basic form the direct approach based on the Brightness Constancy Assumption (BCA) is:

$$\min_{\mathbf{p}} \sum_{\mathbf{q}\in\mathfrak{R}} d(\mathbf{q};\mathbf{p})^2, \tag{1}$$

with $d(\mathbf{q};\mathbf{p}) = \|S(\mathbf{q}) - T(\mathcal{W}(\mathbf{q};\mathbf{p}))\|$, $S$ the source image, and $T$ the target image. An image $I$ is considered as a continuous function from $\Omega_I \subset \mathbb{R}^2$ to $\mathbb{R}^h$ with $h$ the number of colour channels and $\Omega_I$ the domain of the image. Bilinear interpolation is used to evaluate an image at non-integer locations. The cost function in (1) can be made robust so as to handle occlusions and specularities that violate the BCA. A classical framework is to use an M-estimator, as in [OB95, AGKM07]:

$$\min_{\mathbf{p}} \sum_{\mathbf{q}\in\mathfrak{R}} \rho\big(d(\mathbf{q};\mathbf{p})\big). \tag{2}$$

As will be seen later, our contribution exploits the properties of saturated M-estimators (an M-estimator is said to be *saturated* when it is constant above a certain threshold). In this paper, we use Tukey's bisquare $\rho$-function without any loss

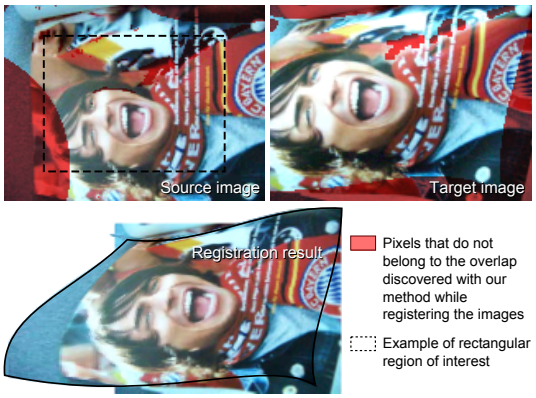


**Figure 1:** *We propose a new algorithm that does not require one to define a region of interest (*RoI*). Our algorithm discovers the exact overlap between two images while registering them. Using the rectangular* RoI *in dashed line defeats classical methods since it contains pixels that do not belong to the overlap.*

of generality:

$$\rho(x) = \begin{cases} \frac{c^2}{6}\left(1-\left(1-\frac{x^2}{c^2}\right)^3\right) & \text{if } |x| \leq c \\ \frac{c^2}{6} & \text{otherwise,} \end{cases} \qquad (3)$$

with $c$ a constant determining the sensitivity to outliers of the M-estimator. A value $x$ such that $\rho(x) = \frac{c^2}{6}$ corresponds to an outlier. Ideally, this constant should be 4.685 times the standard deviation of the noise contained in the images [AGKM07]. In practice, we consider that the standard deviation equals 20% of the maximal pixel value.

Problem (1) is a non-linear least-squares problem that can be solved iteratively using for instance the Gauß-Newton algorithm [Bjö96]. Problem (2) can be solved with an Iteratively Reweighed Least-Squares algorithm [DKSP03].

**The overlap and the RoI.** The direct approach to image registration is interesting because it does not rely on feature correspondences. However, standard registration algorithms require a RoI $\mathfrak{R}$ included in the overlap of the images. This is a difficult 'chicken-and-egg' problem since the overlap can only be determined after a successful registration. There is no known satisfactory solution to this problem.

Let $\mathfrak{O}_S$ be the image *overlap*, *i.e.* the set of pixels of the source image that are also seen in the target image:

$$\mathfrak{O}_S = \left\{ \mathbf{q} \in \Omega_S \mid \underline{\mathbf{q}}' \in \Omega_T \text{ and } S(\mathbf{q}) \approx T(\underline{\mathbf{q}}') \right\}, \qquad (4)$$

where $\underline{\mathbf{q}}'$ is the pixel $\mathbf{q}$ transformed with the true deformation between $S$ and $T$. It is obvious that the cost function in (1) or in (2) cannot be evaluated at those pixels that do not belong to $\mathfrak{O}_S$. As a consequence, $\mathfrak{R}$ must be included in $\mathfrak{O}_S$, otherwise the registration algorithms based on (1) and (2) will fail. Besides, it is better to have a RoI as large as possible in order to have the greatest quantity of information to estimate the warp. The problem here is that the real overlap $\mathfrak{O}_S$ is known only after a successful registration of the images.

**Previous work.** As we review in §2.1, the RoI is often a polygonal region in the source image defined either by the user or by some *ad hoc* means [Bar08]. This lacks automatism and may be unreliable. The adaptive RoI [PA04] is another approach. It considers the entire domain of the source image as an initial RoI and updates it during the optimization process. As we review in §2.2, the cost function of [PA04] is extremely hard to minimize and has global minima that do not correspond to the correct solution (see figure 2).

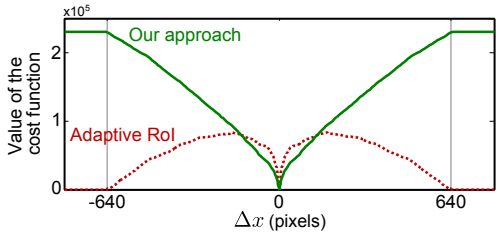


**Figure 2:** *Profile of the cost functions of the adaptive* RoI *approach of [PA04] (red dashed curve) and our approach (green solid curve). The source and the target images are 640 pixels wide. The simulated warp is a translation along the x-axis parametrized by $\Delta x$ (more details in §2.2). The cost function of [PA04] vanishes for a warp that creates no overlap while our cost function has only one global minima that corresponds to the true translation (i.e. $\Delta x = 0$).*

**Contribution.** We propose a novel approach to direct image registration. It is fundamentally different from standard approaches in that it does not need a RoI. This is made possible by considering the off-target pixels as outliers; the theoretical foundations of this principle are explained in §3. The cost function we propose to optimize takes into account *all* the pixels of the source image. A fixed penalty that corresponds to the one given to usual outliers is associated to the off-target pixels. We then use the standard robust M-estimation framework of equation (2) to handle both the usual outliers and the off-target pixels in a unified way. Our new approach has several advantages. First, the proposed cost function does not have trivial minima (see figure 2). Second, it solves all the above-mentioned problems related to the RoI. Third, the overlap is automatically obtained as the set of inliers.

Although generic, our approach is experimented with two specific warps: the homographic and the Free-Form Deformation warps, briefly reviewed in §4.1. The robustness and accuracy of our approach is compared to other ones in §4.2 and §4.3.

## 2. Region of Interest: State of the Art

### 2.1. Rectangular Region of Interest

A common approach used to define the RoI consists in guessing a maximal per-pixel displacement. The RoI is then chosen as a rectangle obtained by removing to the source image domain a margin of width larger than the hypothesized maximal displacement. Ideally, the width of this margin should be as close as possible to the actual maximal displacement, rarely known before registration. The margin width is commonly overestimated so that the optimization algorithm will not fail. Nonetheless, a large RoI provides more information to estimate the warp accurately. Moreover, the size of the RoI affects the profile of the cost function in (1). A simple experiment inspired by [PA04] illustrates this phenomenon. Figure 3 shows, for different margin sizes, the evolution of the cost function *versus* a single shift parameter $\Delta x$ (the amplitude of a translation along the $x$-axis). The source and the target images are identical (shown in figure 4) except for a Gaussian noise with standard deviation equal to 5% of the maximal pixel value. Figure 3 shows that a small margin (a large RoI) results in a smooth cost function but has dramatically restricted range of admissible translations. Using a larger margin (a smaller RoI) increases the range of possible translations but creates a lots of local minima in the cost function.

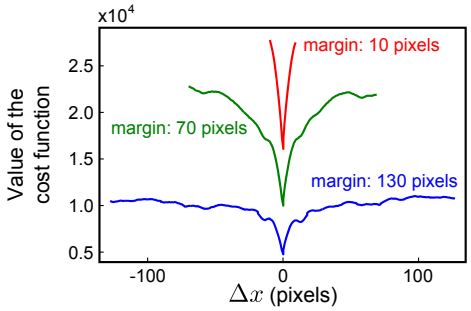


**Figure 3:** *Profile of the cost function in problem (1) for rectangular* RoI *with margins ranging from 10 to 130 pixels (for images of size* $640 \times 480$*).*

| | Large margin (small RoI) | Small margin (large RoI) |
|---|---|---|
| Range of admissible transformations | + | − |
| Quantity of information available for the registration | − | + |

**Table 1:** *Respective advantages and disadvantages of the large and small margins. Note that neither of them has all the advantages.*

### 2.2. Adaptive Region of Interest

An alternative to the rectangular RoI has been proposed in [PA04]. In this approach, the fixed RoI $\mathfrak{R}$ is replaced by an adaptive RoI $\mathfrak{R}_A(\mathbf{p})$:

$$\min_{\mathbf{p}} \sum_{\mathbf{q}\in\mathfrak{R}_A(\mathbf{p})} d(\mathbf{q};\mathbf{p})^2. \tag{5}$$

For a given set of parameters $\mathbf{p}$, $\mathfrak{R}_A(\mathbf{p})$ contains all the pixels (except for a 1-pixel margin used to compute the target image derivatives by finite differences) from the source image that, once warped, belongs to the domain of the target image, *i.e.* $\mathfrak{R}_A(\mathbf{p}) = \{\mathbf{q} \in \Omega_S \mid \mathbf{q}' \in \Omega_T\}$ with $\mathbf{q}' = \mathcal{W}(\mathbf{q};\mathbf{p})$. Although this method does not require one to define a RoI, it is not fully satisfactory. First, problem (5) is badly posed in the sense that there exists an infinite number of minima that do not correspond to the correct warp parameters. These minima appear when there is no overlap between the source and the warped target images. This fact is illustrated with an experiment similar to the one used in §2.1. We observe in figure 2 that the cost function of problem (5) is null (and thus minimal) as soon as the domains do not overlap ($|\Delta x| > 640$). Second, the fact that $\mathfrak{R}_A(\mathbf{p})$ depends on $\mathbf{p}$ makes problem (5) hard to solve rigorously. The authors of [PA04] propose to neglect the dependency on $\mathbf{p}$ and alternate the estimation of $\mathfrak{R}_A$ and $\mathbf{p}$. Third, the adaptive RoI algorithm is not robust to outliers and, as such, it cannot properly handle occlusions and specularities.

## 3. Direct Image Registration without Region of Interest

We propose a new method to direct image registration that does not need a RoI. It thus avoids the above mentioned problems related to the RoI. Our new cost function uses all the pixels of the source image, as the adaptive RoI of [PA04]. However, as the example of figure 3 shows, our cost function has no trivial minima. We will show that it is also much easier to optimize rigorously. The key idea of our method is to penalize the off-target pixels with a fixed cost. The cost associated to the other pixels remains the usual robust colour discrepancy of (2). To some extent, this maximizes the size of the overlap between the two images. We use the same penalty for the off-target pixels and the outlying pixels, for reasons explained below.

**Derivation.** Imagine a target camera with an unbounded field of view. Such a camera would produce images with an infinite domain. Imagine now that a plane with a rectangular hole is placed between the camera and the observed scene, as figure 4(b) illustrates. The part of the scene visible through the hole corresponds to the actual target image $T$. The rest of the scene is not seen because it is *occluded* by the plane, exactly as for the pixels hidden by an external occluder, as shown in figure 4(a). With this reasoning, it becomes natural for one to handle off-target pixels as usual outliers.

A direct yet incomplete mathematical statement of our

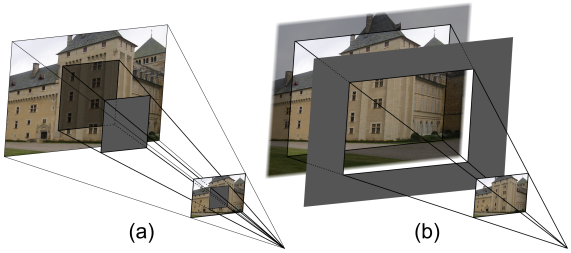


**Figure 4:** *Pixels out of the field of view (b) can be considered as usual outliers (a).*

idea as a minimization problem is:

$$\min_{\mathbf{p}} \sum_{\substack{\mathbf{q}\in\Omega_S \\ \mathbf{q}'\in\Omega_T}} \rho\big(S(\mathbf{q})-T(\mathbf{q}')\big) + \sum_{\substack{\mathbf{q}\in\Omega_S \\ \mathbf{q}'\notin\Omega_T}} \frac{c^2}{6}. \tag{6}$$

Solving (6) is difficult since two sums are mixed, with a number of terms varying as a function of $\mathbf{p}$ since $\mathbf{q}' = \mathcal{W}(\mathbf{q};\mathbf{p})$. First of all, we rewrite the fixed penalty term:

$$\min_{\mathbf{p}} \sum_{\substack{\mathbf{q}\in\Omega_S \\ \mathbf{q}'\in\Omega_T}} \rho\big(S(\mathbf{q})-T(\mathbf{q}')\big) + \sum_{\substack{\mathbf{q}\in\Omega_S \\ \mathbf{q}'\notin\Omega_T}} \rho(x_0), \tag{7}$$

where $x_0$ is any value saturating the M-estimator: $\rho(x_0) = \frac{c^2}{6}$. With the bisquare ρ-function, any value $x_0$ such that $|x_0| \geq c$ is suitable (see equation (3)). Problem (7) can be rewritten:

$$\min_{\mathbf{p}} \sum_{\mathbf{q}\in\Omega_S} \rho\Big([\mathbf{q}'\in\Omega_T]\big(S(\mathbf{q})-T(\mathbf{q}')\big) + [\mathbf{q}'\notin\Omega_T]x_0\Big), \tag{8}$$

where $[\,]$ is the operator such that $[a]=1$ if $a$ is true and $[a]=0$ otherwise. We rewrite (8) by introducing the image $T_\infty$:

$$T_\infty(\mathbf{q}) = \begin{cases} T(\mathbf{q}) & \text{if } \mathbf{q}\in\Omega_T \\ \alpha & \text{otherwise} \end{cases} \quad \text{and} \quad \Omega_{T_\infty} = \mathbb{R}^2, \tag{9}$$

where α is any value such that $\|S(\mathbf{q})-\alpha\| > x_0$. Finally, our method is to solve:

$$\min_{\mathbf{p}} \sum_{\mathbf{q}\in\Omega_S} \rho\big(S(\mathbf{q})-T_\infty(\mathbf{q}')\big). \tag{10}$$

Problem (10) is solved with standard Iteratively Reweighed Least-Squares.

**M-estimator and overlap.** An interesting property of our approach is that it automatically discovers the overlap. For instance, with Tukey's bisquare M-estimator, a pixel $\mathbf{q}$ such that $\big(\big|\rho\big(S(\mathbf{q})-T_\infty(\mathbf{q}')\big)-\frac{c^2}{6}\big| \leq \varepsilon\big)$ can be considered as an outlier (with ε a small constant, *e.g.* $10^{-4}$). The overlap in the source image is the set of source pixels verifying this condition. The overlap in the target image is the warped source overlap. Recovered overlaps are illustrated in figure 1 and in §4.3.

## 4. Experimental Results

### 4.1. Deformation Models

**Homographic warp.** The homography warp $\mathcal{W}_H$ is a global deformation model that explains the relationship between two images taken by a camera that either rotates around its optical centre or observes a planar scene. With $\mathbf{p}\in\mathbb{R}^8$ and $\mathbf{q}=(x,y)$, it is defined by:

$$\mathcal{W}_H(\mathbf{q};\mathbf{p}) = \frac{1}{p_7x+p_8y+1}\begin{pmatrix} p_1x+p_2y+p_3 \\ p_4x+p_5y+p_6 \end{pmatrix}.$$

**Free-Form Deformation warp.** The B-spline warp is a particular type of Free-Form Deformations [RSH*99]. It is parametrized by a set of 2D control points $\mathbf{p}_{ij}$ acting as 'attractors' for the deformation. The control points $\mathbf{p}_{ij}$ ; $i\in\{1,\dots,n_x\}$, $j\in\{1,\dots,n_y\}$ are grouped in a vector $\mathbf{p}\in\mathbb{R}^{2n_xn_y}$. The analytical form $\mathcal{W}_B$ of this warp is the 2D tensor product of the 1D cubic B-spline:

$$\mathcal{W}_B(\mathbf{q};\mathbf{p}) = \sum_{j=1}^{n_y}\sum_{i=1}^{n_x} \mathbf{p}_{ij}N_i(x)N_j(y),$$

where $N_i$ is the $i$th cubic B-spline basis function [dB01].

### 4.2. Synthetic Data

**Data generation.** We generated synthetic data in the following manner. First, a warp (homography or B-spline) is determined by interpolating some randomly generated point correspondences. The source image is obtained by unravelling a texture image with the previously computed warp and the texture image is used as the target image. The average distance between the point correspondences controls the warp magnitude γ (in pixels). A proportion α of the source and target images is then replaced with data from a different image to simulate occlusions. Last, Gaussian noise with standard deviation σ is added to the images. We used colour images with intensities coded with real values between 0 and 1. The images are $320\times240$ pixels wide. Figure 5 gives an illustration of the generation process.

**Experimental setup.** The influence of several factors is studied: the transformation magnitude γ, the amount of noise σ and the proportion of erroneous data α. Each one of these factors is studied independently with default values: $\gamma=8$ pixels, $\alpha=10\%$ and $\sigma=0.1$ (10% of the maximal pixel intensity value). Several algorithms are compared: rectangular ROI (RECT), the adaptive ROI of [PA04] (ADAP) and our approach (MAXC). Different variants of the RECT algorithm are considered: narrow (10%) and large (25%) margins without M-estimator (RECTN, RECTL) and with M-estimator (RECTNM, RECTLM). The reported results are averages over 100 trials.

**Optimization failures.** As explained in §2.1, a ROI of fixed size can lead to a failure of the optimization process. Figure 6 shows in which proportion such failures occur for the

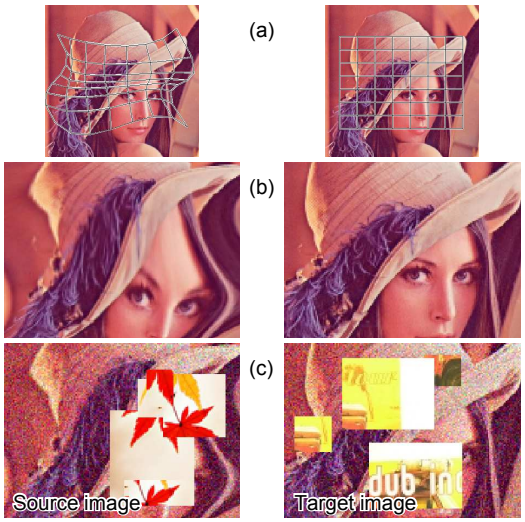


**Figure 5:** *Synthetic data generation. (a) Texture image and deformation used to generate the source and the target images. (b) The warp is unravelled to generate the source image. (c) Noise and outliers are added to the images.*

experiments of the next 3 paragraphs and for the default values. Note that convergence towards a false solution (local minimum) is not counted as a failure. We observe that there are more failures with a wide rectangular ROI (RECTN) than with a small one (RECTL). There are less failures with an M-estimator (RECTNM, RECTLM) than without because the steps of the optimization algorithms tend to be smaller. In the sequel, when an algorithm fails to converge, the reported measurements are from the last valid iteration.

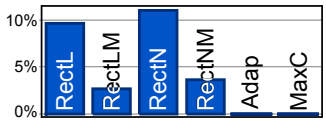


**Figure 6:** *Failure rates.* ADAP *and* MAXC *never fail because they do not rely on a fixed* ROI.

**Number of iterations.** Figure 7 shows the number of iterations. Overall, the convergence is faster with the homographic than with the B-spline warp. This comes from the fact that the homographic warp is global. The apparent rapidity of the algorithms relying on a rectangular ROI stems from the fact that these algorithms can fail before convergence when the given ROI is not valid. Our approach, MAXC is generally better than ADAP which is the only other method that does not require a ROI. However, MAXC takes more iterations to converge when the transformation magnitude is large. This is explained by the fact that many pixels from the source image, once transformed, do not belong to the target image domain. The convergence is slightly slowed down since these pixels are penalized with our approach.

**Geometric error.** Figure 8 shows the geometric error, the discrepancy in pixels between the estimated and the ground truth transformations. We observe that the amount of noise does not influence much the performance of the algorithms. On the contrary, the geometric error is influenced by the transformation magnitude and by the proportion of outliers. This is especially true for the approaches that do not include an M-estimator. Compared to the other methods, our approach is the one that gives the best results. We can see that, with our approach, the geometric error is often less than one pixel. This result is particularly important because it shows that our approach is not biased by the penalty term used for the pixels which are warped outside of the target domain.

**Photometric error.** The average photometric error obtained after the last iteration of the studied algorithms is reported in Figure 9. The smallest errors are always obtained with our approach whatever the varying factor and the geometric transformation are.

### 4.3. Real Data

**Overlap.** We consider a source and a target images of a planar scene taken from two different view points and with an occlusion in the target image. Under these conditions, the warp between the two images is a homography. Figure 10 shows the ROI used during the last iteration of four different algorithms. This ROI is shown in both the source and the target images. The difference image between the warped target and the source images is also shown. It shows that our approach, MAXC, is the only one to estimate the correct homography. The main point of figure 10 is that the final ROI determined with MAXC corresponds exactly to the true overlap between the images. The ROI used by ADAP at convergence does not take into account the occluder. Consequently, ADAP is not able to recover correctly the homography. The ROI utilized by RECTLM does not contain enough pixels making this approach unable to determine precisely the homography. Finally, the algorithm RECTNM fails to converge since its ROI contains pixels that do not belong to the overlap (figure 10 shows the last valid iteration).

**The widest panorama.** We consider a video captured by a camera that rotates around its optical centre with a uniform movement from left to right. Consequently, the successive images are linked with homographies. The goal of this experiment is to build a panorama as wide as possible by taking the first image of the video and the furthest image for which the registration is successful. As shown in figure 11, the widest panoramas are obtained with ADAP and MAXC. For this video, there are no occluders and, thus, the results of ADAP and MAXC are similar. The algorithms RECTN and RECTL get the smallest panoramas since the maximal displacements are dictated by margin sizes.

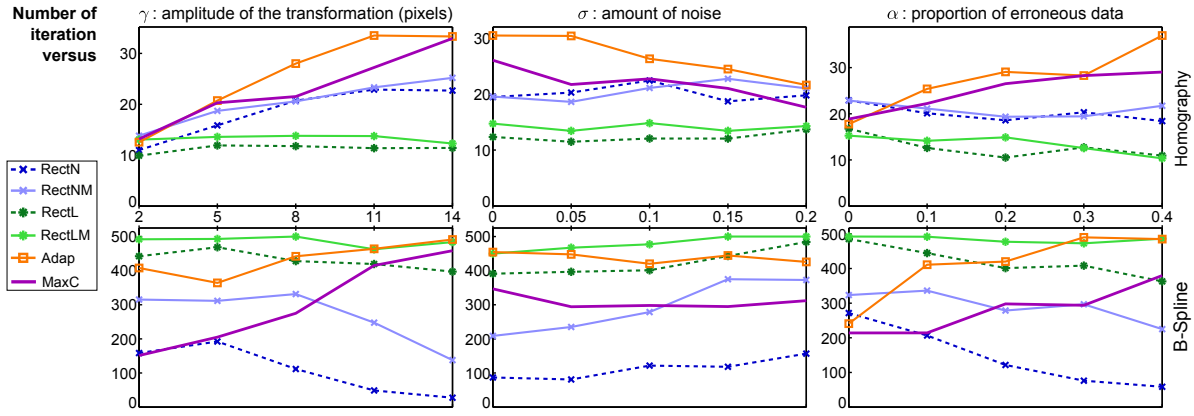


**Figure 7:** *Influence of several factors on the the number of iterations. The number of iterations done by the algorithms based on a rectangular* ROI *is relatively low because these methods can stop prematurely (fail) as soon as the* ROI *is not valid.*

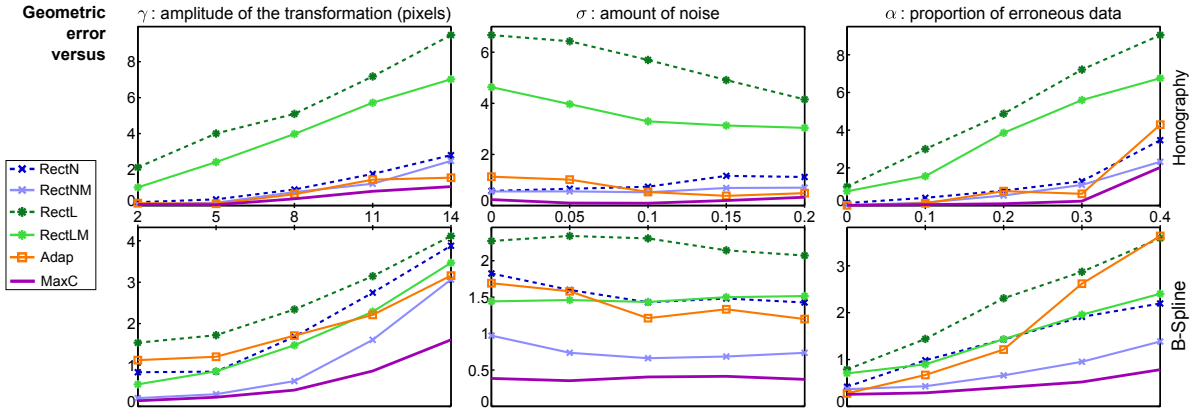


**Figure 8:** *Influence of several factors on the geometric error. Our approach (*MAXC*) gives the best results. Globally, the approaches relying on M-estimators are the best ones.*

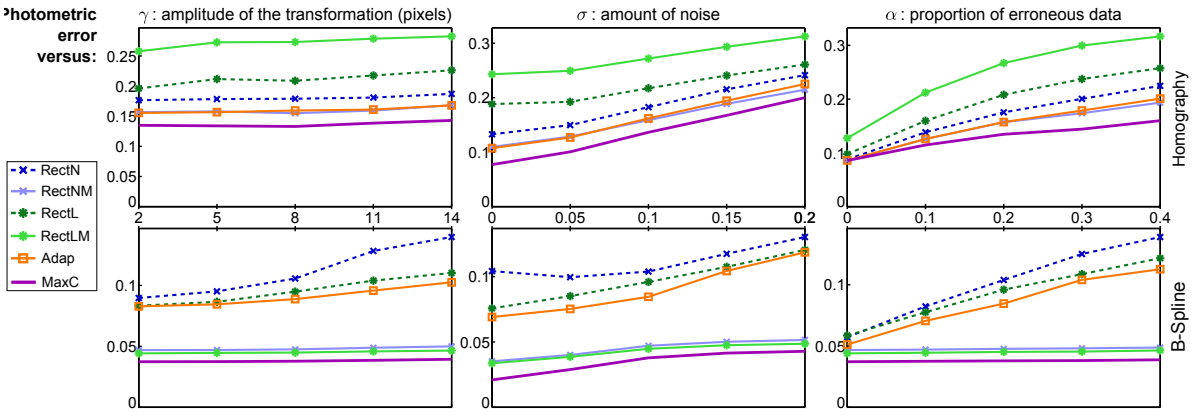


**Figure 9:** *Influence of several factors on the photometric error. The best results are always obtained with our approach whatever the transformation model and whatever the studied factor.*

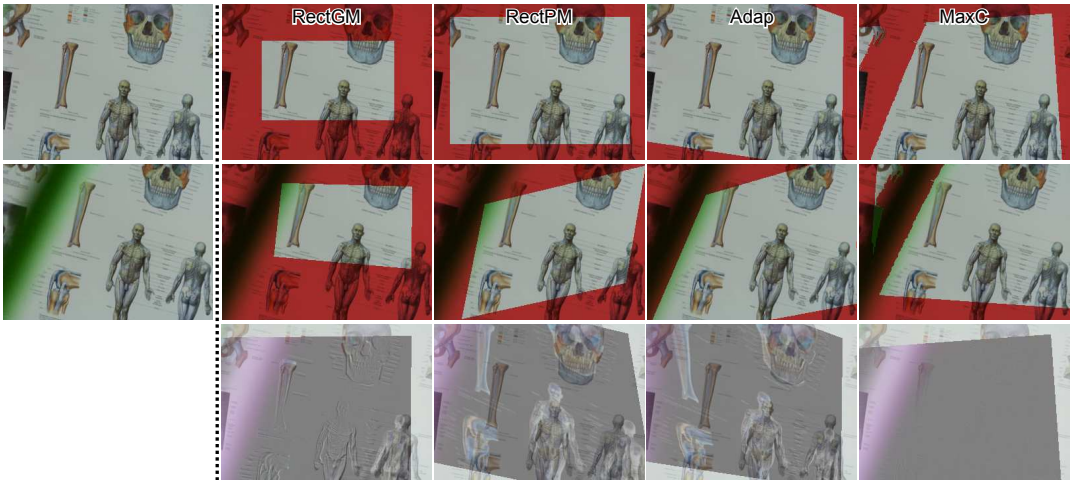


**Figure 10:** *Examples of registration results for different algorithms. The first row corresponds to the source image, the second row to the target image, and the last row to the difference between the source and the warped target image. The red pixels are the pixels not included in the* ROI *during the last iteration of the algorithms. Note that the* ROI *computed by our approach (*MAXC*) corresponds to the true overlap (taking into account both the field of view and the occluder). Our approach is the only one that successfully registers this pair of images.*

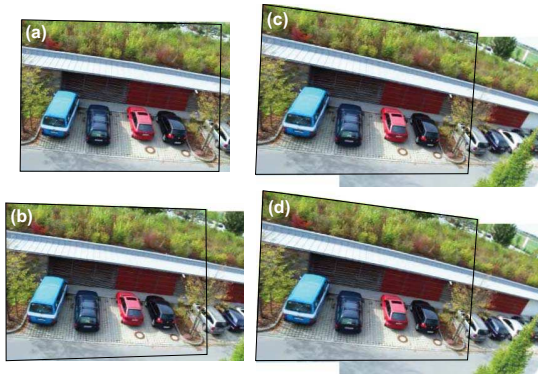


**Figure 11:** *Panorama calculated with (a)* RECTN*, (b)* RECTL*, (c)* ADAP *and (d)* MAXC*. The widest panoramas are obtained with* ADAP *and our approach:* MAXC*.*

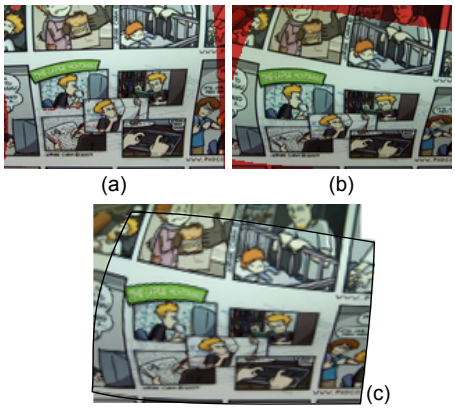


**Figure 12:** *Example of deformable mosaic. (a): source image ; (b): target image ; (c) mosaic. The red pixels in (a) and (b) are the pixels that do not belong to the overlap determined with our approach.*

**Deformable mosaic.** An example of deformable registration using our method is given in figure 12. This figure illustrates that our approach automatically retrieves the true overlap in both the source and the target images. Note that a video corresponding to that example is provided as supplemental material.

**Pattern tracking.** Figure 13 illustrates the tracking of a pattern in a video sequence. Three approaches are compared: our approach, and two approaches using a fixed rectangular ROI (defined with either a large margin or a narrow margin). The object to track is a deforming banknote. We thus use a FFD warp with $5 \times 5$ control points. The pattern (*i.e.* the source image) to track is defined as a part of the first image in the video sequence. The pattern is registered in each new image (which plays the role of the target image) using as an initial solution the registration determined for the previous image. Figure 13 shows that the approaches relying on fixed ROI fail as soon as a part of the pattern is not visible in the target image. Such problems cannot happen with our approach. Figure 13 also illustrates that, with our approach, the true overlap is correctly determined in both the source (pattern) and the target images. The fourth and fifth columns of figure 13 shows that our approach handles erroneous data (occlusions and specularities) and the overlap in a unified manner.


**Figure 13:** *Pattern tracking in a video sequence. Only a few frames of the video are shown here (the complete video is available as supplemental material). For our method (first and second rows), we systematically show the pattern (i.e. the source image) in order to illustrate the automatic discovery of the true overlap. For the methods that rely on a fixed rectangular* ROI *(third and fourth rows), the pattern is shown only once since it does not vary with time. The approaches relying on a fixed* ROI *fails prematurely because some pixels of the* ROI *are warped outside of the target image domain (frame #300 with a large margin and #708 with a narrow margin). The frames #867 and #1096 shows how our approach handles occlusions and specularities.*

## 5. Conclusion

We proposed a new approach to image registration that does not need a ROI. It relies on a theoretical foundation stating that it is possible to consider the off-target pixels as outliers. This new point of view of direct image registration resulted in a slight but elegant modification of the cost function usually optimized. An interesting consequence of our approach is that the true overlap between the images is simply the set of inlying pixels. Compared to previous approaches, ours solves the problems related to the ROI and to the optimization of the cost function. The efficiency of our approach was illustrated with extensive experiments. In particular, we showed that our approach was better than the previous methods in term of accuracy and robustness.

## References

[AGKM07] ARYA K. V., GUPTA P., KALRA P. K., MITRA P.: Image registration using robust M-estimators. *Pattern Recognition Letters 28* (2007), 1957–1968. 1, 2

[Bar08] BARTOLI A.: Groupwise geometric and photometric direct image registration. *IEEE Transactions on Pattern Analysis and Machine Intelligence 30* (2008), 2098–2108. 1, 2

[BGM04] BAKER S., GROSS R., MATTHEWS I.: Lucas-Kanade 20 years on: A unifying framework. *International Journal of Computer Vision 56* (2004), 221–255. 1

[Bjö96] BJÖRCK Å.: *Numerical Methods for Least Squares Problems*. SIAM, 1996. 2

[dB01] DE BOOR C.: *A Practical Guide to Splines – Revised Edition*. Springer, 2001. 4

[DKSP03] DEANS M., KUNZ C., SARGENT R., PEDERSEN L.: Terrain model registration for single cycle instrument placement. *Intelligent Robots and Systems* (2003). 2

[HBA97] HARDIE R., BARNARD K., ARMSTRONG E.: Joint MAP registration and high-resolution image estimation using a sequence of undersampled images. *IEEE Transactions on Image Processing 12* (1997), 1621–1633. 1

[HPN99] HEIGL B., PAULUS D., NIEMANN H.: Tracking points in sequences of color images. *Pattern Recognition and Image Understanding* (1999). 1

[IA99] IRANI M., ANANDAN P.: About direct methods. In *Workshop on Vision Algorithms* (1999). 1

[OB95] ODOBEZ J.-M., BOUTHEMY P.: Robust multiresolution estimation of parametric motion models. *Journal of Visual Communication and Image Representation, 6(4):348-365, December 1995. 6* (1995), 348–365. 1

[PA04] PIRES B., AGUIAR P.: Registration of images with small overlap. *IEEE Workshop on Multimedia Signal Processing* (2004). 2, 3, 4

[RSH*99] RUECKERT D., SONODA L., HAYES I., HILL D., LEACH M., HAWKES D.: Nonrigid registration using Free-Form Deformations: Application to breast MR images. *IEEE Transactions on Medical Imaging 18* (1999), 712–721. 4

[Sze06] SZELISKI R.: Image alignment and stitching: A tutorial. *Foundations and Trends in Computer Graphics and Vision 2* (2006), 1–104. 1

[TZ99] TORR P. H. S., ZISSERMAN A.: Feature based methods for structure and motion estimation. *Vision Algorithms* (1999). 1