

RESEARCH

Open Access

Absolute joint moments: a novel image similarity measure

Hrvoje Kalinić^{1*}, Sven Lončarić¹ and Bart Bijmens²

Abstract

In this paper, we propose a novel approach for estimating image similarity. This measure is of importance in assessing image correspondence or image alignment and plays an important role in image registration. Currently, this problem is approached rather one-dimensionally since most registration methods consider the problem as either mono- or multi-modal. This perspective leads to the selection of some form of either the correlation coefficient (CC) or mutual information (MI) as image similarity measure (ISM). We propose a more generic framework for ISM construction, based on absolute joint moments, which can be considered as a generalization of CC. Within this framework, we propose a specific ISM that provides a different trade-off between MI and CC in terms of performance and computational cost for general registration problems. To illustrate this, we compared CC and MI with the proposed ISM and performed extensive experiments with regard to accuracy, robustness and speed. The evaluation demonstrated that the proposed absolute joint moments is a good combination of properties of CC and MI, with respect to speed and performance.

1 Introduction

Image registration is an optimization process that utilizes a similarity measure to find the optimal alignment of two images. The registration accuracy depends on the selection of the optimization algorithm and geometric transformation as well as on the definition of the similarity measure. Therefore, it is essential to use a suitable similarity measure for a given problem. The selection of an image similarity measure, especially in the case of medical image registration, is usually reduced down to the question whether a multi- or mono-modal registration is required. This black-and-white perspective leads to well-known answers and results in the selection of correlation coefficient (CC) for mono-modal registration and mutual information (MI) for multi-modal registration.

However, the problem of registration can be approached from several perspectives, and often, the registration of the same two images can be performed using different methodological approaches. For example, Zitova et al. in [1] distinguish not only a multi-modal approach but

also a multi-view or multi-temporal approach. Similarly, Maintz et al. in [2], classify registration not only from a intra-/inter-modality perspective but also from a intra-/inter-subject perspective. Perhaps for this reason, many other image similarity measure (ISM) emerged over time (see Section 2).

In the following sections, we propose a wider framework for the construction of image similarity measures. Within this framework, we propose a specific ISM to show that, when constructed in this way, it combines beneficial properties of the two most used ones: CC and MI.

Whatever similarity measure is used for image registration, it has to satisfy one basic condition - at the exact location of the correct alignment of two images, the similarity (measure) has to be maximal. To find the maximum, the image registration incorporates an optimization algorithm, which iteratively calculates a similarity measure. The number of calculations may thus easily reach a number of a few thousands, especially if the geometric transformation is non-linear. Therefore, the complexity of the similarity measure also plays an important role in registration methods. For this reason, in our experiments, we measure the performance of the similarity measure in terms of accuracy, robustness and speed.

*Correspondence: hrvoje.kalinic@fer.hr

¹Department of Electronic Systems and Information Processing, Faculty of Electrical Engineering and Computing, University of Zagreb, Zagreb 10000, Croatia

Full list of author information is available at the end of the article

2 Background

The purpose of an ISM is to quantify the similarity between two images, usually referred to as source and target images. In this section, we will briefly discuss the properties of the two most prominent ISMs in order to compare them to the newly proposed one.

To set the notation, let us denote source and target images as $S(\mathbf{x})$ and $T(\mathbf{x})$, where $\mathbf{x} \in \mathbb{R}^n$ stands for the pixel coordinate vector. Further, the average value is denoted with a line above the function (e.g. \bar{S}), and p is used to denote the (joint) probability density function (e.g. $p_{TS}(\mathbf{x}, \mathbf{y})$).

It is well known [3,4] that CC, defined as

$$CC(T, S) = \frac{\sum_{\mathbf{x} \in \mathbb{R}^n} (T(\mathbf{x}) - \bar{T})(S(\mathbf{x}) - \bar{S})}{\sqrt{\sum_{\mathbf{x} \in \mathbb{R}^n} (T(\mathbf{x}) - \bar{T})^2 \sum_{\mathbf{x} \in \mathbb{R}^n} (S(\mathbf{x}) - \bar{S})^2}}, \quad (1)$$

is an optimal choice if the pixel values between the images $S(\mathbf{x})$ and $T(\mathbf{x})$ are related by an affine function, even in the presence of a reasonable amount of noise. Since most mono-modal image registration problems assume this type of functional relationship between images, CC was considered a dedicated approach for this type of problems and thus is often the first choice.

Similarly, MI defined as

$$MI(T, S) = \sum_{\mathbf{x}, \mathbf{y} \in \mathbb{R}^n} p_{TS}(\mathbf{x}, \mathbf{y}) \log \frac{p_{TS}(\mathbf{x}, \mathbf{y})}{p_T(\mathbf{x})p_S(\mathbf{y})} \quad (2)$$

is established as the appropriate ISM for multi-modal registration [5,6].

All good properties, as well as the shortcomings, of MI and CC come directly from their definition. For example, CC is fast since it uses only summary statistics. The calculation of mean and standard deviation only, which are sufficient to describe an affine functional relationship between pixel values, makes CC a good and fast measure for registration of images with affine relationship between their pixel values. Many image registration problems violate the assumption of an affine relationship between image pixel values, thus making CC a suboptimal choice for images with more complex functional relationship between pixel values [3,7]. On the other hand, the definition of MI includes much more statistical information about the relationship of the images $S(\mathbf{x})$ and $T(\mathbf{x})$. This property comes directly from the joint probability density function $p_{TS} = p(T, S)$, which fully characterizes the relationship between images T and S , without any assumptions on a type of functional relationship or any noise that may exist in the image acquisition process. Many of the disadvantages of the MI come from the very same things that boost its good properties - the joint probability density function (PDF). For example, since

the joint PDF needs to be estimated/approximated, various problems can emerge, like sensitivity to sample size, number of histogram bins or interpolation [7,8].

In this perspective, we can say that the main difference between CC and MI is whether the ISM incorporates (few) summary statistic values or all the statistical information in the form of a PDF. From this perspective, we can thus see that many measures, proposed earlier in the literature, are derivatives of these two most prominent ones. Roughly, we can classify them as either CC-based (e.g. [7,9-16]), or MI-based (e.g. [17-23]), which are sometimes also referred as information theory-based similarity measures (see [24-27]).

For this reason, we will primarily focus on CC and MI, investigate their properties and compare them to the ISM that we propose. In the following section, we will give some motivating examples to show that there is room for improvement besides the existing ISMs. Next, we aim to propose a framework for ISM construction which will utilize a chosen amount of statistical values instead of only a few (such as standard deviation, skewness, kurtosis). In this way, we aim to present a framework which will bridge a gap between CC and MI and select a measure from this framework to show that it incorporates the good sides of both of them. In a way, this paper can be seen as an extension of the idea presented in the paper by Kalinić et al. [28].

3 Motivation

In previous sections, we mentioned that CC is considered ideal for mono-modal image registration due to its elegance and swiftness and with the implicit assumption that the relationship between pixel intensity values is affine and only corrupted by Gaussian noise in mono-modal image acquisition [3,4]. Due to its intrinsic properties, MI can be used in both mono-modal and multi-modal applications, the latter of which is extensively used. This section provides three simple examples where either CC or MI (or both) do not perform well (see Figure 1).

As will be shown in further sections, the performance of the MI may vary, depending on the number of bins selected to approximate the PDF. To distinguish one MI implementation from another, the number of bins is used as index (e.g. MI_8 and MI_{256}). The three examples shown in Figure 1 are given for each ISM implementation: CC, MI_8 and MI_{256} . The comparison between the images is done while applying three different geometric transformations: scaling, rotation and translation. The images which are to be aligned are constructed from the same template image selected from the test set (see Section 5.1) by simple noise addition, with $A_n/A_s = 0.5$ (for details about noise degradation model, see Section 5.2). Since both images are constructed from the same template, the correct alignment is already known, i.e. the ISM should have a maximal value for unity scaling, zero rotation and zero translation.

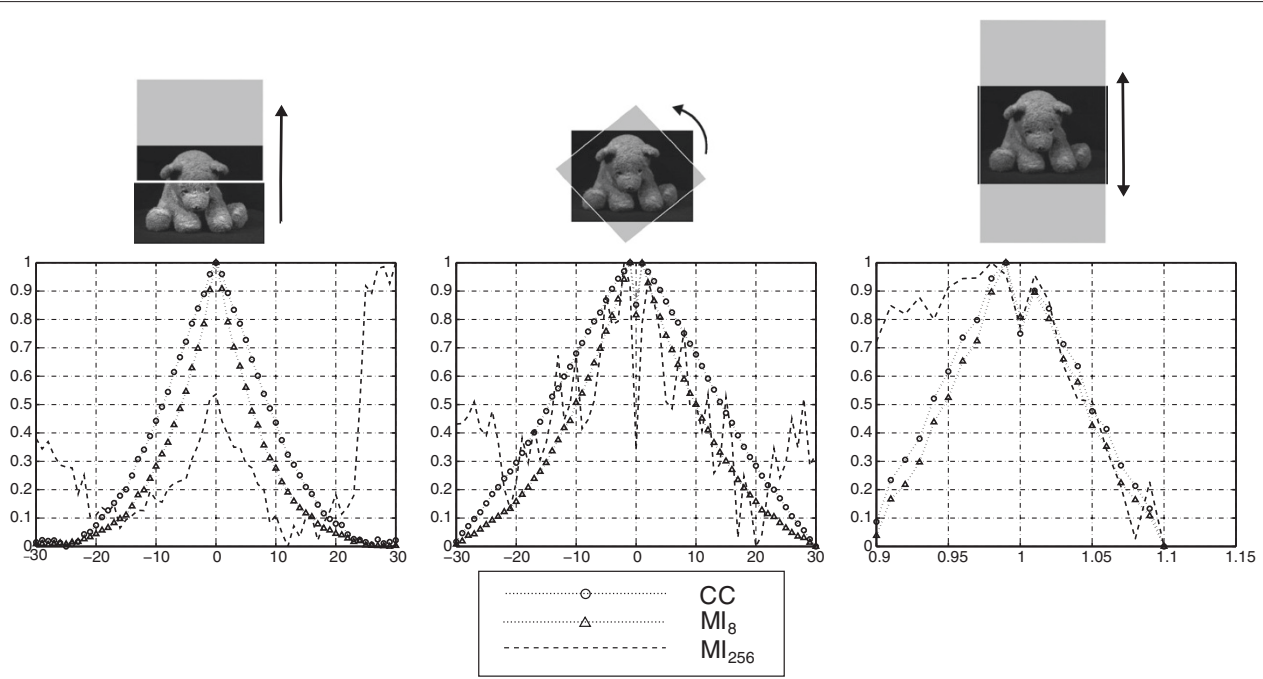


Figure 1 Graphs showing the behaviour of CC, MI_8 and MI_{256} for translation, scaling and rotation of an image.

In the first example, the image is consecutively translated by 1 pixel in the range of -30 to 30 . In the second example, the image is consecutively rotated by 1° in the range of -30° to 30° . In the third example, the image is scaled by $s = 0.9 + 0.01 \cdot n$ where $n \in \{0, \dots, 20\}$. Figure 1 shows the values for MI_8 , MI_{256} and CC calculated for the examples.

Notice that MI_{256} does not perform well in the first example, and neither of the ISMs has the maximum at the correct location for both rotation and scaling. Therefore, an ISM defined in a different way might be able to produce a better result in these examples, but then, it obviously remains to be seen if its better performance would be related only to this particular image. Both questions will be addressed further on (see Section 7).

4 Absolute joint moments

In this section, we define source and target images as random variables and denote them as S and T instead of $S(\mathbf{x})$ and $T(\mathbf{x})$. The following notation is used: if $\mu_S = E[S]$, then $E[(S - \mu_S)^n]$ stands for the n th central moment of the random variable S , or more generally, $E[(S - \mu_S)^n(T - \mu_T)^m]$ stands for the joint central moment of the order (m, n) of the random variables S and T .

4.1 Framework for constructing image similarity measures

As mentioned in the previous section, the idea is to use more statistical information from the images (instead of

only a limited amount as in CC) for the ISM construction. Joint moments are the proposed information to use. The motivation behind this becomes clear if we write Equation 1 in the following form:

$$CC(T, S) = \frac{E[(T(\mathbf{x}) - \mu_T)(S(\mathbf{x}) - \mu_S)]}{\sigma_T \cdot \sigma_O}, \quad (3)$$

where σ_T and σ_S denote standard deviations.

If we rewrite Equation 3 to the following form

$$CC(T, S) \cdot \sigma_T \cdot \sigma_O = E[(T(\mathbf{x}) - \mu_T)(S(\mathbf{x}) - \mu_S)], \quad (4)$$

we can see that this is equal to the joint central moment of order $(1, 1)$. From this, we propose a generalization in the following form:

$$AJM = \left| \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \frac{1}{\omega_n} \frac{1}{\omega_m} \cdot E[(T - \mu_T)^n(S - \mu_S)^m] \right|. \quad (5)$$

We refer to the absolute joint moment (AJM) framework for ISM construction. Notice that if we take only the first element of the sum and select $\omega_n = \omega_m = 1$, this reduces down to the absolute covariance (numerator of CC from Equation 3).

4.2 Proposed image similarity measure

For the proposed image similarity, we use a specific selection of weights ω_n and ω_m , which are computationally efficient and guarantee the convergence of the sum. Since this is just one of many possible selection of weights, we will denote ISM as AJM_i to indicate that is just one instance.

Using the selected weights, Equation 5 can be rearranged into the form of the Taylor expansion of the following exponential function:

$$\begin{aligned} AJM_i &= \left| \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \frac{1}{n!} \frac{1}{m!} E[(T - \mu_T)^n (S - \mu_S)^m] \right| = \\ &= \left| E \left[\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \frac{1}{n!} \frac{1}{m!} (T - \mu_T)^n (S - \mu_S)^m \right] \right| = \\ &= \left| E \left[\sum_{n=1}^{\infty} \frac{1}{n!} (T - \mu_T)^n \sum_{m=1}^{\infty} \frac{1}{m!} (S - \mu_S)^m \right] \right| = \\ &= \left| E \left[(e^{S - \mu_S} - 1)(e^{T - \mu_T} - 1) \right] \right|. \end{aligned} \quad (6)$$

As the expectation is estimated on the overlapping region of images S and T (here denoted by D_X), we will only have the estimation of AJM_i :

$$\widetilde{AJM}_i = \frac{1}{N} \left| \sum_{x \in D_X} (e^{S - \mu_S} - 1)(e^{T - \mu_T} - 1) \right| \quad (7)$$

where N denotes the number of pixel pairs in D_X . In further sections, the \widetilde{AJM}_i will be the only used and tested ISM from this framework, so we will denote it simply as AJM .

We will show that this ISM selected from the proposed framework, compared to MI and CC, will have a different trade-off between speed and performance as will result directly from the definition of AJM . In order to show this, in further sections, we will investigate the properties of

AJM and compare them to MI and CC with regard to robustness, accuracy and speed.

5 Experimental data

5.1 Data set

Images from publicly available databases were used to test the properties of the registration implemented using AJM as similarity measure. The test set was constructed so as to have as much diversity as possible. First, we used all 44 miscellaneous images from the SIPI database [29]. Since this database does not have medical images, all 19 medical images from the VIS database [30] were added to the set as well as 3 mammography images from the MIAS database [31]. From the MIAS set, only three images were selected since the variability of the images from this set is low. Finally, 34 images of different objects from the ALOI database [32] were added to the set which made the total number of images in the testing set 100. The set constructed in this way contained images with different context, from natural to artificially constructed images. Both colour and greyscale images were represented in this set, but all images were converted to greyscale before the registration. All the images were coded with either 7 or 8 bits per pixel and the resolution of images ranged from 128×128 to $1,024 \times 1,024$. For the purpose of the paper, all images were converted to double-precision floating-point format and scaled to interval $[0, 1]$.

A few images from each database, forming a representative collection of images from the test set, are depicted in the first row of Figure 2. In the second row of Figure 2, the same images after artificial degradation is shown (see Section 5.2 for details). The pairs of images constructed in this way will be used for the evaluation of the ISMs in most of further experiments.

5.2 Image degradation model

To evaluate the performance of the ISMs, we will need a pair of each image. Therefore, we introduced several



Figure 2 A representative collection of the images used in our experiments.

degradation models that will be applied on each image in order to simulate different effects that may happen during the image acquisition process such as excessive noise, contrast changes or non-linear intensity distortion. The image degradation models are described in the following subsection and, as can be noticed, are inspired by the paper of Maes et al. [33].

5.2.1 Contrast inhomogeneity

A linear spatially variant intensity transformation is used to simulate image contrast inhomogeneity effects and is modeled by the following expression:

$$I'(x, y) = \alpha(x, y)I(x, y), \quad (8)$$

where $\alpha(x, y)$ is defined as follows:

$$\alpha(x, y) = \frac{1}{1 + k_1 \cdot ((x - x_c)^2 + (y - y_c)^2)}, \quad (9)$$

with (x_c, y_c) being the coordinates of the point around which the curve is positioned and k_1 is the distortion parameter.

5.2.2 Noise

To assure that noise does not randomly affect the contrast, we selected a distribution with finite support. Thus, the uniform distribution seemed a natural choice as a noise degradation model. Notice that this simple degradation model becomes more complex when the sequential image degradation model is applied since that degradation affects the noise distribution as well. Additive uniform noise from the interval $[0, k_2 A_s]$ is superimposed on the original image. Here, A_s stands for the amplitude of the signal and k_2 stands for the amplitude ratios between noise and signal (A_n/A_s).

5.2.3 Non-linear intensity distortion

A non-linear intensity transformation is used to simulate pixel value distortion and is described by the following polynomial:

$$f(I_{xy}) = (I_{xy} - i_1)(I_{xy} - i_2)(I_{xy} - i_3) \dots (I_{xy} - i_n), \quad (10)$$

where I_{xy} stands for the intensity level (at position (x, y)) and $i_1, i_2, i_3, \dots, i_n$ are roots of the polynomial that simulates the intensity distortion. After each distortion, the image is normalized to keep the original range of pixel values.

6 Experiments and results

In the first two experiments (Sections 6.1 and 6.2), the image pairs, between which the correct alignment is to be determined, are the original and the degraded image. Therefore, the gold standard is well known since the correct alignment is for unity scaling, and zero translation and rotation. To evaluate the performance of the ISM, an exhaustive search for the global maximum was done in order to assure that the suitability of the ISM, rather

Table 1 Parameter range for the robustness test

Parameter	Range
(x_c, y_c)	Image center
k_1	$[0.0001, 0.0004]$
k_2	$[0.1, 1]$
i_1	$[0.1, 0.5]$
i_2	0.5
i_3	$[0.5, 0.9]$
$i_1 + i_3$	1

than the search strategy, is evaluated. Each ISM between image pairs is calculated for a progressive shift of 1 pixel in the interval $[-100, 100]$, for stepwise rotation of 1° in the interval $[-180, 179]$ and for scaling in steps of 0.01 to increase or decrease the scaling factor in the interval $[0.5, 2]$.

6.1 Robustness test

Various image degradations may alter the intensities and may affect the performance of a similarity measure. To evaluate the robustness of a similarity measure with respect to additive noise, contrast change or non-linear intensity distortion, each degradation model (described in Section 5.2) was applied to each image from the set (see Section 5.1), and the similarity between the original and the degraded image was calculated for different translation, rotation and scaling factor. For the degradation model, the parameters listed in the Table 1 were used.

As can be assumed from the Table 1, only third-order polynomials were used for the robustness test. However, this selection ensures non-linear distortion, as can be seen in Figure 3.

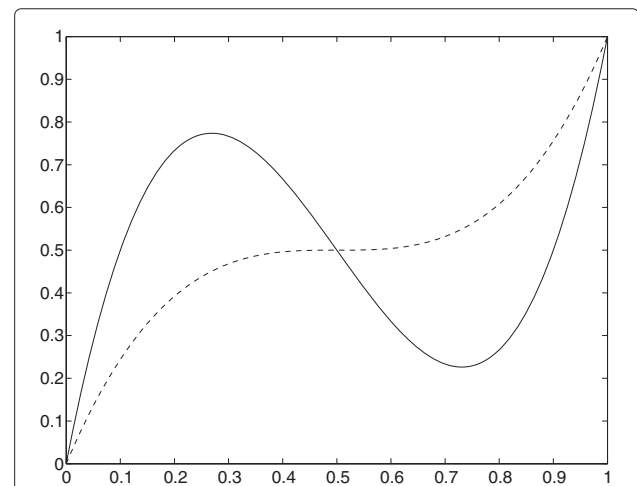


Figure 3 Polynomials used in the robustness test. The solid line represents the case $i_1 = 0.1$, while the dashed line represents the case $i_1 = 0.5$.

Here, we will focus on the robustness of the ISMs with respect to the amount of deformation. If the ISM is robust to the degradation effects, the maximum of the ISM will be achieved for approximately the same transformation for any amount and type of degradation. Therefore, in this test, we assess how the number of images with alignment error less than ξ changes with respect to the amount of degradation. ξ is defined as 5 pixels for translation, 0.05 for scaling (see Equation 11 for error definition), and 9° for rotation - which represents 5% of the Figure 4 x -axis range (maximal deformation). The experiment was done for each degradation model, and the alignment error was measured for three different types of image transformation: translation, scaling and rotation. Only 1 degree of freedom was allowed in each degradation model, namely

k_1 , k_2 and i_1 . Each parameter was changed in ten equidistant steps within the range allowed by Table 1. The experimental results are given in Figure 5. Graphs in Figure 4 are sorted so that each row shows results for a different image degradation, and each column shows results for a different geometric transformation used.

As anticipated, this experiment showed that MI_{256} is not robust to noise and that CC is not robust to non-linear intensity distortions. For AJM, one can observe that it is affected by high non-linear intensity distortion. However, for a moderate distortion, it still performs satisfactorily. Therefore, AJM seems robust to noise and to moderate amounts of non-linear intensity distortion. AJM is also fairly robust to contrast inhomogeneity since the performance of AJM is comparable to those of other ISMs. All

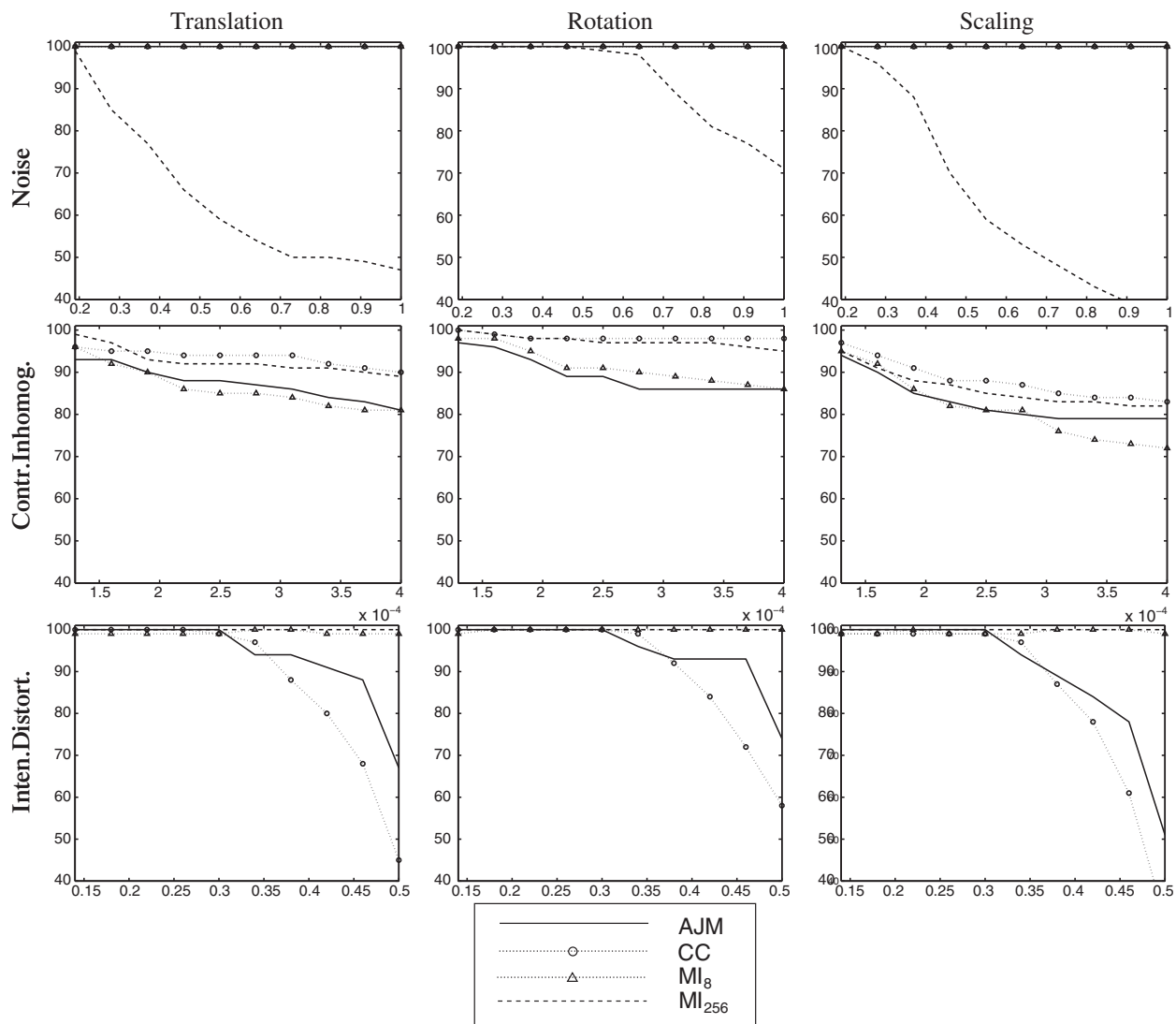


Figure 4 Distributions of the ISM errors. The graphs show the total number (y-axis) of aligned images with an error lower than x (value on the x-axis). The performance of CC, AJM, MI_8 and MI_{256} is plotted using circles, solid line, triangles and dashed line, respectively.

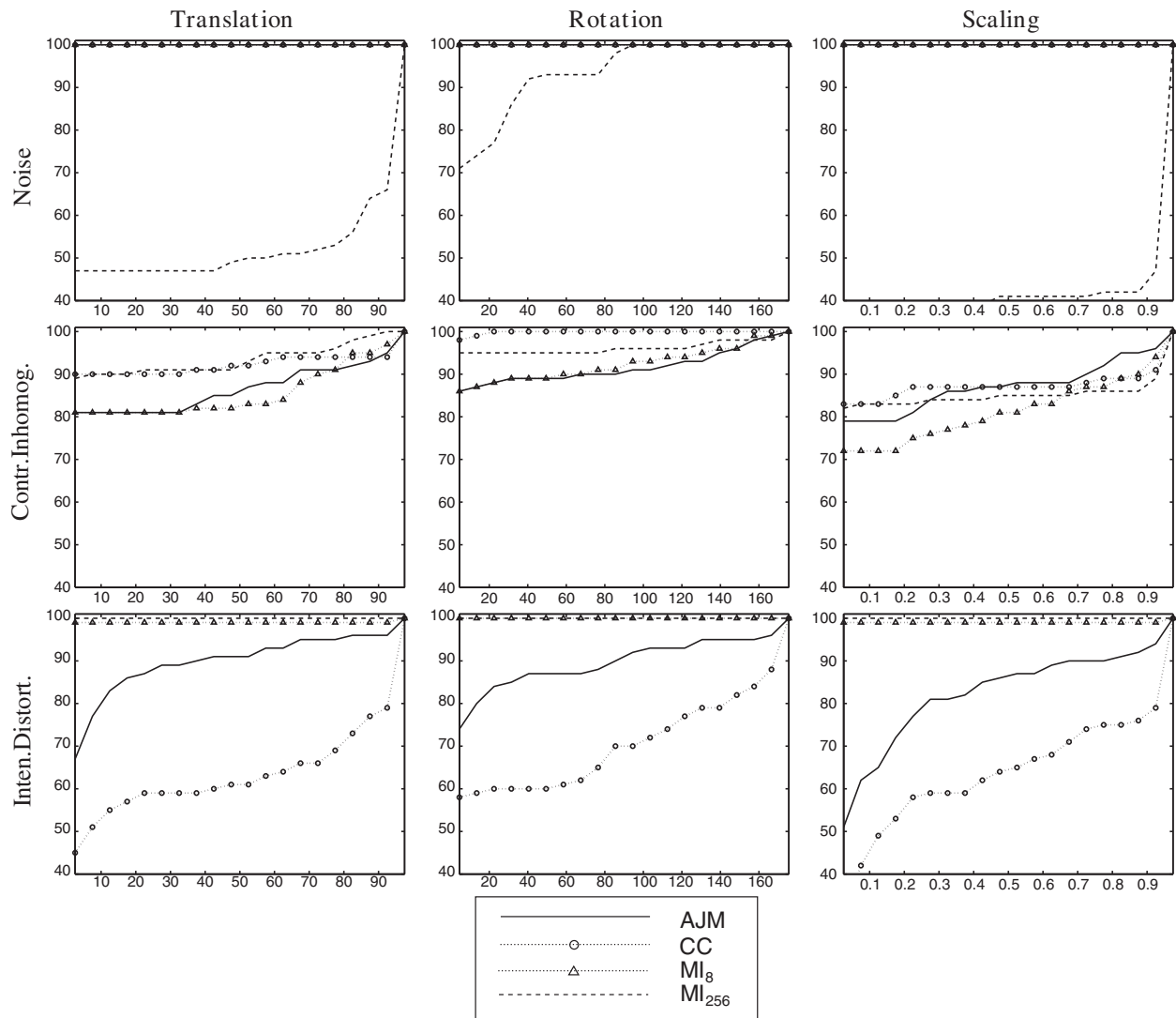


Figure 5 Robustness of ISMs. The graphs show change in the performance of each ISM with regard to distortion. For contrast inhomogeneities, the parameter is k_1 ; for noise, k_2 ; and for non-linear distortion, i_1 . For each ISM and for each distortion, the y-axis shows the the number of images with alignment error less than ξ and the x-axis shows the amount of distortion (i.e. k_1 , k_2 and i_1). The performance of CC, AJM, MI₈ and MI₂₅₆ are plotted using circles, solid line, triangles and dashed line, respectively.

of these conclusions hold for translation, rotation and scaling.

The results of the previous experiment could vary if a different error threshold ξ is selected. To estimate how this results would change for different ξ selected, we performed the following test. Again, all three degradation models were applied and the images were aligned after three different transformations. However, the fixed parameters in this case are k_1 , k_2 and i_1 . The parameters are set to represent the largest degradation. The results are presented in Figure 5, where, in each line of the table, a graph of the ISMs for different distortion is plotted, and where, in each column, a different transformation is used to achieve the correct image alignment. The

graphs show how many images are aligned with an error lower than a certain amount and therefore give insight in the distribution of the ISM error for the image data set.

Figure 5 shows that the order of the ISMs for their relative performance would remain approximately the same no matter what error threshold (ξ) is selected. For all experiments (with the exception of the combination of rotation and contrast inhomogeneities), AJM is between CC and MI with regard to the overall number of correct alignments and sometimes even between the two different implementations of MI. This was, as might be assumed, from its theoretical properties.

As expected, the experiments showed that no ISM could compete with the performance of MI for large non-linear intensity distortions. However, the use of higher order moments was helpful since AJM performs better than CC for non-linear intensity distortion. From the graphs, it is also clear that in a noiseless environment, MI₂₅₆ performs better than MI₈, but in a noisy environment, MI₂₅₆ is not such a good choice.

6.2 Accuracy test

This test is used to evaluate the accuracy of localizing the exact position of the maximum of the ISM. The image pairs for the experiments are the original images that form the test set and the degraded images, both of which are explained in Section 5. The degraded images are constructed by applying all three degradation models, where first additive noise is added, then contrast degradation is simulated, and finally, non-linear intensity distortion is done on the pixel value. For degradation of each image, the parameters are selected randomly within the value range listed in the Table 2.

As can be noticed in Table 2, the range of parameters is a bit larger in this test than in the previous one. Thus, higher contrast inhomogeneity is allowed as well as lower noise; similarly, non-linear image distortion is implemented as an n th order polynomial, where n can range from 2 to 6. Notice that, according to the robustness test, the increase of the range of these values will not go in favour of AJM.

The alignment error is measured separately for translation, scaling and rotation and the results are summarized in Table 3. First column shows the average absolute error calculated for translation, and the average absolute errors for rotation and scaling are given in second and third columns, respectively.

The error for translation and rotation is given in pixel and degrees, and the error for scaling is a unitless value that is calculated as follows:

$$\epsilon_S = \log_2(D_i), \quad (11)$$

where D_i is the scaling factor (deformation amount) for which the maximal ISM value is achieved. Index i stands for the measurement (image) number, and N denotes the total number of measurements (images). This is done so

Table 3 The average absolute error for translation, scaling and rotation

	$\overline{\epsilon_T}$	$\overline{\epsilon_R}$	$\overline{\epsilon_S}$
CC	51.27	72.17	0.49
AJM	19.57	31.04	0.24
MI ₈	15.48	23.70	0.15
MI ₂₅₆	33.01	37.73	0.45

that the scaling error is symmetrical, i.e. it gives the same error for squeezing and stretching the image by the same factor. Also, it gives no error if the images are scaled by the same factor.

6.3 Speed test

Since CC, MI and AJM are implemented as defined in Equations 1, 2 and 6, respectively, we can notice that the computational complexity of all three measures is $\mathcal{O}(N)$. However, it is expected that CC will work faster than AJM since it has to calculate only a product and a ratio instead of a product of exponential functions. Similarly, we can expect that AJM is faster than MI since it does not require the histogram formation. To evaluate this, we measured the execution time of all three algorithms. For this purpose, the Matlab profiler was used.

The algorithm was implemented on a standard quad-core PC without parallelization. The computation time required for one evaluation of the similarity measure varies linearly with the number of samples in the overlapping region between the two images. The results are shown in Table 4. The speed performance evaluation is calculated as the average from 1,000 function calls for an image size of 512×512 .

7 Overall comparison

At the beginning of the paper (Section 3), some examples were given, where CC, MI₈ and MI₂₅₆ are unable to find the exact alignment. To investigate how AJM performs in a similar setting, the same experiments were repeated using AJM. The results are presented in Figure 6, showing the performance of AJM alongside CC, MI₈, or MI₂₅₆.

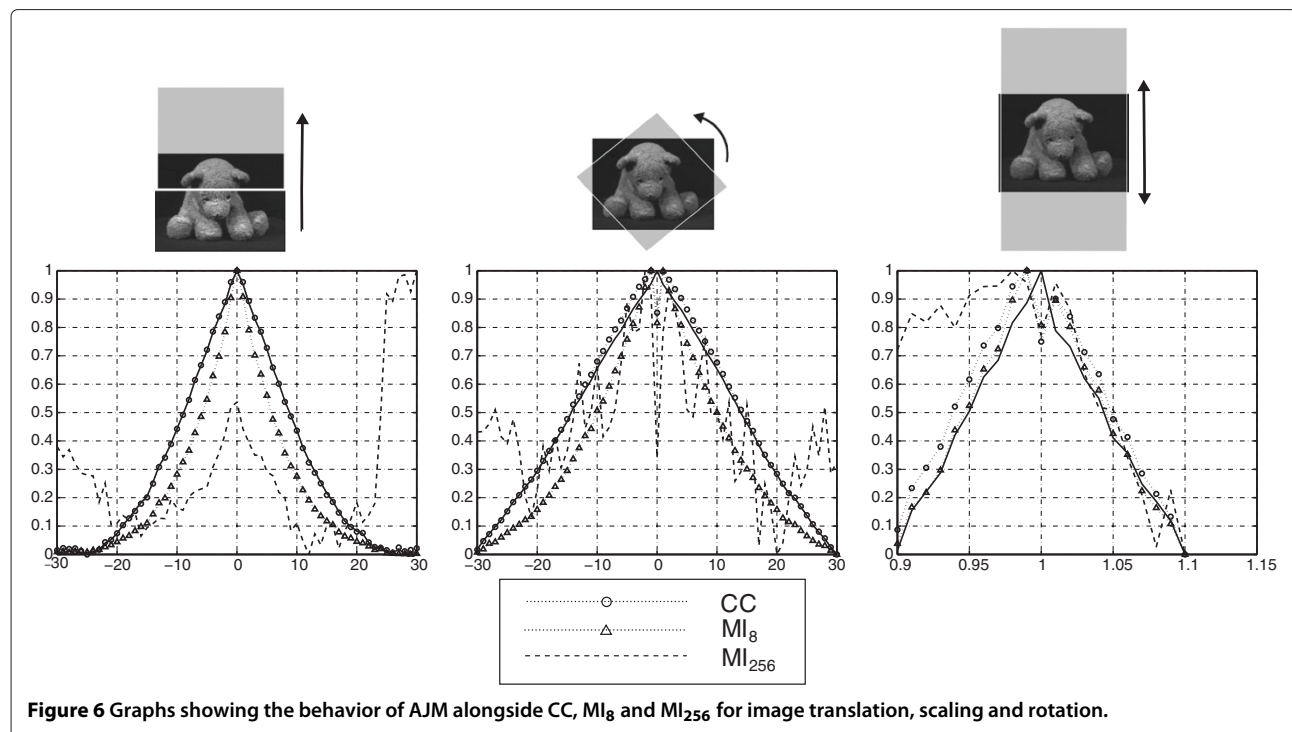
Table 2 Parameter range for the accuracy test

Parameter	Range
(x_c, y_c)	Within image
k_1	[0.00005, 0.0005]
k_2	[0, 1]
$i_1 \dots i_n$	[0, 1]
n	{2, 3, 4, 5, 6}

Table 4 Execution time measurements

	Time
CC	20.772 ms
AJM	24.609 ms
MI ₈	50.010 ms
MI ₂₅₆	78.627 ms

The performance of each ISM is evaluated as average performance of 1000 function calls.



As can be noticed, in all three cases, AJM outperforms the other ISMs. To check that this is due to the nature of this particular image, we ran this experiment on all images from the data set. To recall, both images were created from the same original image by simple noise addition, and we tried to align them using translation, rotation and scaling. In each experiment, we counted the number of cases where AJM outperforms the other ISMs. The results are given in Table 5.

8 Conclusions

The experiments show that AJM is robust to noise, fairly robust to contrast inhomogeneities, more robust than CC and less robust than MI for non-linear intensity distortion. The robustness test also showed that MI_{256} is not very robust to noise. Overall, we can say that AJM is positioned between CC and MI in terms of robustness.

The accuracy test shows that AJM is less accurate than MI_8 and overall better than MI_{256} . This difference

between different MI implementations also emphasizes how MI is affected by the number of bins of the histogram and the interpolation techniques, while AJM is not. Both MI and AJM outperform CC significantly, primarily due to the fact that CC cannot cope with non-linear intensity distortion. An additional strength of AJM is that it does not require a statistically significant number of pixels for the calculation of entropy, so it can be calculated for a smaller region compared to MI.

Although all three ISMs have similar complexity, AJM is a faster method than MI since it does not require histogram calculations (nor estimations), but it is still slower than CC. The comparison is qualitatively summarized in Table 6.

As a general conclusion, we can say that the experiments have shown that the proposed ISM is able to determine the correspondence among images with complex relationships between the pixel values and is computationally

Table 5 The total number of aligned image pairs for which AJM outperforms other ISMs

	CC	MI_8	MI_{256}
Translation	1	0	66
Rotation	69	60	67
Scaling	68	51	72

The evaluation was done with respect to different geometric transformations. Image pairs were altered using a noise degradation model. Since the total number of images in the data set is 100, the numbers in the table also reflect the percentage of image pairs for which AJM outperforms other ISMs.

Table 6 Performance of similarity measures

	CC	MI	AJM
Noise	+	-/+	++
Contrast inhomogeneity	+	+	+
Non-linear intensity distortion	-	++	+
Speed	++	-	+
Insensitive to number of bins	+	-	+
Insensitive to interpolation	+	-	+

The table qualitatively summarizes the properties of all three ISMs.

more efficient and does not have some of the inherent disadvantages of MI. Therefore AJM_i , from the AJM framework for ISM construction, is complementary to CC and histogram-based approximation of MI, and the specific application will guide the optimal ISM choice.

Competing interests

The authors declare that they have no competing interests.

Authors' information

HK received his PhD degree in Electrical Engineering from the University of Zagreb, Croatia. He is a research and teaching assistant at the University of Zagreb, Faculty of Electrical Engineering and Computing and participates in national and international research projects as a member of the Image Processing Group at the Department of Electronic Systems and Information Processing. His main field of interest is digital image processing and analysis. He is the author of five abstracts, seven conference papers and four journal papers. He is a program committee member or reviewer of four scientific conferences and one international journal.

SL received his BSc and MSc degrees in Electrical Engineering from the University of Zagreb, Zagreb, Croatia in 1985 and 1989, respectively, and his PhD degree in Electrical Engineering from the University of Cincinnati, Cincinnati, OH in 1994. From 1990 to 1994, he was a Fulbright Fellow and a Research Assistant with the University of Cincinnati. He was an assistant professor at the New Jersey Institute of Technology, Newark from 2001 to 2003. He has been a full professor at the University of Zagreb since 2006, where he was an assistant professor from 1996 to 2001 and an associate professor from 2001 to 2006. He was the Chair of the Department of Electronic Systems and Information Processing, Faculty of Electrical Engineering and Computing, University of Zagreb. He is a coauthor of more than 180 refereed journal and conference publications. He served as the Editor-in-Chief of the *Journal of Computing and Information Technology*. SL served as the IEEE Croatia Section Chair from 2005 to 2008. He currently serves as the chapter's Chair for IEEE Signal Processing Society.

BB is an ICREA research professor at the Department of Information and Communication Technologies of the Universitat Pompeu Fabra, Barcelona. While he graduated with a master's degree in Electro-Mechanical Engineering Sciences, he obtained a PhD in Medical Sciences on advanced echocardiographic data acquisition and analysis for addressing clinical and pathophysiological questions. He became an associate professor at the Medical Faculty of the KU Leuven where he established a Cardiac Imaging Research Group. He extended his experience in a mixed research/hospital management position in St. George's Hospital in London (UK) and initiated a new interdisciplinary Cardiology-Engineering research group in Zagreb (Croatia). Currently, he is performing multidisciplinary research in cardiovascular pathophysiology and image analysis/modelling, in collaboration with the major hospitals in Barcelona. His research interests are translational cardiovascular pathophysiology, focussing on assessing cardiac function and understanding and recognising the changes induced by disease and how treatment strategies can be used to modulate this remodelling. BB is the author of more than 130 peer-reviewed papers in international journals (both in medicine and engineering), more than 70 full proceedings papers in international conferences (mainly in biomedical engineering), over 360 abstracts at international conferences (mainly Medical) and more than 130 invited lectures.

Author details

¹Department of Electronic Systems and Information Processing, Faculty of Electrical Engineering and Computing, University of Zagreb, Zagreb 10000, Croatia. ²ICREA, Universitat Pompeu Fabra, Barcelona E08018, Spain.

Acknowledgements

This study was supported by the Ministry of Science, Education and Sports of the Republic of Croatia under grant 036-0362214-1989 any partially by the Subprograma de Proyectos de Investigacin en Salud, Instituto de Salud Carlos III, Spain (FIS - PI11/01709).

Received: 22 November 2012 Accepted: 3 April 2013

Published: 26 April 2013

References

1. B Zitova, J Flusser, Image registration methods: a survey. *Image Vis. Comput.* **21**(11), 977–1000 (2003)
2. J Maintz, M Viergever, A survey of medical image registration. *Med. Image Anal.* **2**, 1–36 (1998)
3. M Sonka, J M Fitzpatrick (Eds), *Handbook of Medical Imaging, Volume 2. Medical Image, Processing and Analysis*. (SPIE Press, 2000)
4. P Viola, Alignment by maximization of mutual information, PhDthesis, MIT (1995). [Co-supervisors: Lozano-Perez, Tomas and Atkeson, Christopher G]
5. A Collignon, F Maes, D Delaere, D Vandermeulen, P Suetens, G Marchal, in *Information Processing in Medical Imaging*, ed. by Y Bizais, C Barillot, and R Di Paola. Automated multimodality image registration based on information theory (Kluwer Dordrecht, 1995), pp. 263–274
6. P Viola, WM Wells, Alignment by maximization of mutual information. *Int. J. Comput. Vis.* **24**, 137–154 (1997)
7. A Roche, G Malandain, X Pennec, N Ayache, The correlation ratio as a new similarity measure for multimodal image registration, in *MICCAI '98*, ed. by WM Wells, A Colchester, S Delp. Proceedings of the First International Conference on Medical Image Computing and Computer-Assisted Intervention. Lecture notes in computer science, vol. 1946 (Springer-Verlag, London, 1998), pp.1115–1124
8. JPW Pluim, JBA Maintz, MA Viergever, Interpolation artefacts in mutual information-based image registration. *Comput. Vis. Image Underst.* **77**(9), 211–232 (2000)
9. KV Arya, P Gupta, PK Kalra, P Mitra, Image registration using robust M-estimators. *Pattern Recogn. Lett.* **28**, 1957–1968 (2007)
10. A Fitch, A Kadyrov, Christmas W, J Kittler, Fast robust correlation. *IEEE Trans. Image Process.* **14**(8), 1063–1073 (2005)
11. H Foroosh, JB Zerubia, M Berthod, Extension of phase correlation to subpixel registration. *IEEE Trans. Image Process.* **11**(3), 188–200 (2002)
12. S Kaneko, Y Satoh, S Igarashi, Using selective correlation coefficient for robust image registration. *Pattern Recognit.* **36**(5), 1165–1173 (2003)
13. Y Keller, A Averbuch, A projection-based extension to phase correlation image alignment. *Signal Process.* **87**, 124–133 (2007)
14. S Periaswamy, H Farid, Medical image registration with partial data. *Med. Image Anal.* **10**(3), 452–464 (2006)
15. H Yan, JG Liu, in *Proceedings of the British Machine Vision Conference*. Robust phase correlation based motion estimation and its applications (Leeds, September 2008), pp. 1–4
16. AJH Hui, CE Hann, JG Chase, EEW Van Houten, Fast normalized cross correlation for motion tracking using basis functions. *Comput. Methods Prog. Biomed.* **82**, 144–156 (2006)
17. JBA Maintz, EHW Meijering, MA Viergever, in *Image Processing*. General Multimodal Elastic Registration Based on Mutual Information (SPIE, New York, 1998), pp. 144–154
18. H Luan, F Qi, Z Xue, L Chen, D Shen, Multimodality image registration by maximization of quantitative-qualitative measure of mutual information. *Pattern Recognit.* **41**, 285–298 (2008)
19. Y Wang, BG Hu, Derivations of normalized mutual information in binary classifications. *Fuzzy Syst. Knowl. Discov. Fourth Int. Conf.* **1**, 155–163 (2009)
20. ND Cahill, JA Schnabel, JA Noble, DJ Hawkes, in *Medical Imaging 2009:Image Processing*, vol. 7259, ed. by JPW Pluim, BM Dawant. Overlap invariance of cumulative residual entropy measures for multimodal image alignment (SPIE, New York, 2009)
21. D Loeckx, P Slagmolen, F Maes, D Vandermeulen, P Suetens, Nonrigid image registration using conditional mutual information. *Med Imaging. IEEE Trans.* **29**, 19–29 (2010)
22. P Thevenaz, M Unser, Optimization of mutual information for multiresolution image registration. *Image Process., IEEE Trans.* **9**(12), 2083–2099 (2000)
23. SK Warfield, J Rexilius, PS Huppi, TE Inder, EG Miller, III WMW, GP Zientara, FA Jolesz, R Kikinis, A binary entropy measure to assess nonrigid registration algorithms. in *MICCAI '01*. Proceedings of the 4th International Conference on Medical Image Computing and Computer-Assisted Intervention (Springer-Verlag, London, 2001), pp. 266–274
24. WR Crum, DLG Hill, DJ Hawkes, Information theoretic similarity measures in non-rigid registration. *Inf Process, Med Imaging.* **18**, 378–387 (2003)
25. J Zachary, SS Iyengar, Information theoretic similarity measures for content based image retrieval. *J. Am. Soc. Inf. Sci. Technol.* **52**, 856–857 (2001)

26. L Cazzanti, M Gupta, Information-theoretic and set-theoretic similarity. IEEE International Symposium on Information Theory, Seattle. July 2006 (IEEE, New York, 2006) pp. 1836–1840
27. D Lin, An information-theoretic definition of similarity. Proceedings of the 15th International Conference on Machine Learning (Morgan Kaufmann, San Francisco, 1998), pp. 296304
28. H Kalinić, S Lončarić, B Bijenens, A novel image similarity measure for image registration. 7th International Symposium on Image and Signal Processing and Analysis, Dubrovnik, September 2011, ed. by D S vS Lončarić, G Ramponi (IEEE, New York, 2011), pp. 195199
29. SIPI database. <http://sipi.usc.edu/database>. Accessed 22 April 2013
30. VIS database. <http://vis-www.cs.umass.edu/~vislib/Medical/InfarctScan/images.html>. Accessed 22 April 2013
31. MIAS database. <http://peipa.essex.ac.uk/info/mias/>. Accessed 22 April 2013
32. ALOI database. http://staff.science.uva.nl/~mark/aloi/aloi_grey_red4_view.tar. Accessed 22 April 2013
33. F Maes, A Collignon, D Vandermeulen, G Marchal, P Suetens, Multimodality image registration by maximization of mutual information. *IEEE Trans. Med. Imaging.* **16**, 187–198 (1997)

doi:10.1186/1687-5281-2013-24

Cite this article as: Kalinić et al.: Absolute joint moments: a novel image similarity measure. *EURASIP Journal on Image and Video Processing* 2013 **2013**:24.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com