MINIMAL INFORMATION EXCHANGE FOR IMAGE REGISTRATION

Marco Dalai and Riccardo Leonardi

Department of Electronics for Automation,
University of Brescia, via Branze 38, Brescia, Italy.
{name.surname}@ing.unibs.it

ABSTRACT

In this paper we consider the problem of estimating the relative shift, scale and rotation between two images X and Y that are available to two users, respectively A and B, connected through a channel. User A is asked to send B some specifically selected minimal description of image X that will allow B to recover the relative shift, rotation and scale between X and Y. The approach is based on a distributed encoding technique applied to the Discrete Fourier Transform phase and to the Fourier-Mellin transform of the images.

1. INTRODUCTION

In this paper we consider a problem of identification of the relative shift, rotation and scale between two images with minimal information exchange. The problem can be formulated in the following way. Two images X and Y are obtained by cropping a common scene, supposed to be planar, with possible relative shift, rotation and scale difference between the two images. Suppose that X and Y are not available at the same point but are instead available to two users A and B that can communicate across a channel. Suppose user B is interested in knowing the relative shift, rotation scale between the two images. User A is asked to send to B some compact description of image X so as to let him recover the rotation, scale and shift parameters. It is clear that, for example, if user A sends the whole lossless description of the X image, then B can recover those parameters. It is however meaningful to consider that a lossy description of X could also be sufficient for recovering them with good confidence. In general, it is interesting to study the most efficient way to extract a code from the X image for the proposed problem, i.e., to find the coding strategy that allows to reduce as much as possible the number of bits sent from A to B.

The motivations for the definition and the study of this problem, are to be found in different applications. An important example is the case of Distribute Video Coding (DVC), which is the application of Distributed Source Coding principles (DSC) to video coding problems. DSC is a branch of information theory which has been receiving increasing attention from the the publication of [1] by Slepian and Wolf. The main topic considered in this field is the possibility of separately encode correlated sources with the same encoding efficiency as if they were encoded jointly, provided one single decoder has access to the information communicated by every encoder. The application of this principle to the case of Video Coding leads to interesting problems that rely on the exploitation of the correlation between images that are available to remote users (see for example [2]). In this context, an underlying fundamental problem of interest is to find a way to "align" two images in a remote fashion, i.e., without having both images available in one point, but allowing for the
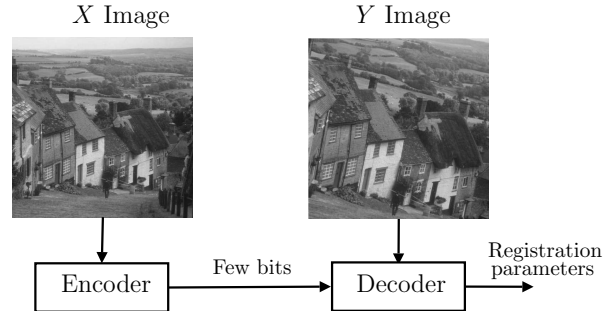

Figure 1: Graphical representation of the considered problem.

communication of a small amount of information. Thus, for example, it is of interest to study as a starting problem the scenario considered in this paper, where an image Y is completely available to user B (say the decoder), which wants to perform the registration, and an image X is available to user A (say the encoder) which has to send to B a concise description of X that brings enough information to perform the compensation.

As far as we know, an ad hoc study of this problem has not been presented in the literature. In the DVC community, more complicated situations are usually considered where motion or disparity compensation are performed at the decoder. In those cases it is usually assumed that a low pass or high pass lossy description of the frame should be sent from encoder to decoder to provide the necessary information needed to estimate the motion or the disparity [3, 4, 5]. This assumption may be considered reasonable, but it is not supported by any theoretical background, and it is not even derived as a heuristic extension of an optimal solution known for a simpler problem. The aim of this paper is precisely to consider the simpler problem of distributed registration and to show how this problem can be approached building upon a theoretical setting. We believe that this is a necessary step for the study of more complicated problems.

A framework for the remote compensation of shifts between images was proposed in [6]; in this paper we propose a more general study for the distributed registration of images that differ not only for a shift, but also for rotation and scale. This problem is decomposed in several steps using a Discrete Fourier Transform (DFT) representation and the so called Fourier-Mellin approach, already known as a classical approach to the problem of image registration (see for example [7]). In particular, it in this paper, the problems of distributed coding of rotation and scale are reduced to the shift problem previously studied in [6] by proper transform

operations. The technique developed for the shift coding can thus be applied for the distributed compensation of such geometric transformations.

The paper is structured in the following way. In Section 2 the distributed encoding technique of shift information presented in [6] is briefly summarized, while in Section 3 the problem of rotation and scale registration is studied. An example of experimental simulation is described in detail in Section 4. Conclusions are future perspectives are finally drawn in Section 5.

2. DISTRIBUTED CODING OF SHIFTS

2.1 Simple 1-D case with circular shift

Before considering the shift between two images, it is useful to study the simpler problem of a circular shift between two 1-dimensional signals. In this section, for an integer m, '$\{\cdot\}_m$' indicates the modulo-m operation.

Suppose we have two N-point signals $X(\cdot)$ and $Y(\cdot)$ which differ only by a circular shift s, with $0 \leq s < S$, $S < N$, i.e.:

$$X(n) = Y(\{n-s\}_N), \quad n = 0, 1, \ldots, N-1. \quad (1)$$

For the sake of simplicity, we only consider the case when both N and S are powers of 2. Let X be known to user A and Y be known to user B. User A has to send a minimum number of bits to B that allow to detect the shift between X and Y. Let us first quantify what we mean by "minimum number of bits". Suppose s is a random variable uniformly distributed between 0 and $S-1$. If X and Y were both known to A, then A would only need to send $\log(S)$ bits to B in order to communicate the value of s. Consider instead the case when A only knows X and B only knows Y. Supported by distributed source coding theory [1], one may wonder whether it is still possible for A to "send" the value of s using only $\log(S)$ bits. Note that "send" is not really appropriate a term, as A could not know the value of s in this case. What we ask is if it is possible for A to send $\log(S)$ bits that will allow B to infer the value of s. We now prove that this is indeed possible by providing a constructive solution to the problem.

Let $\hat{X}(\cdot)$ and $\hat{Y}(\cdot)$ be the DFT of X and Y respectively, and, for every k, let $\Phi_{\hat{X}}(k)$ and $\Phi_{\hat{Y}}(k)$ be the phase of the coefficient $\hat{X}(k)$ and $\hat{Y}(k)$ respectively. From the shift hypothesis on X and Y, the phases of the DFT are related by the following equation

$$\Phi_{\hat{X}}(k) \overset{2\pi}{=} -\frac{2\pi s k}{N} + \Phi_{\hat{Y}}(k). \quad (2)$$

where $\overset{2\pi}{=}$ stands for a modulo-2π congruence. Let us consider the phase of DFT coefficients taken at exponentially spaced positions, i.e.

$$\Phi_{\hat{X}}(N/2), \Phi_{\hat{X}}(N/4), \Phi_{\hat{X}}(N/8), \ldots, \Phi_{\hat{X}}(N/S). \quad (3)$$

We show that a 1-bit quantization of the above phases - for a total amount of $\log(S)$ bits - suffices to recover the value of s at the decoder.

Let us write the binary representation of s as $s_q s_{q-1} \cdots s_1 s_0$, $s_i \in \{0,1\}, i = 0, \ldots, q$. First consider the $N/2$-th DFT coefficient. For this coefficient eq. (2) becomes

$$\Phi_{\hat{X}}(N/2) \overset{2\pi}{=} -\pi s + \Phi_{\hat{Y}}(N/2) \quad (4)$$

$$\overset{2\pi}{=} -\pi s_0 + \Phi_{\hat{Y}}(N/2). \quad (5)$$

It is clear that when $\Phi_{\hat{Y}}(N/2)$ is known, the sign of $\Phi_{\hat{X}}(N/2)$ uniquely determines the value of s_0. So, one bit extracted from $\Phi_{\hat{X}}(N/2)$ (i.e., the sign) allows B to determine the least significant bit of s. Now, by using an iterated procedure, we show by induction that B can reconstruct the binary representation of s using the signs of the considered coefficients. Supposing that, using the signs of $\Phi_{\hat{X}}(N/2), \Phi_{\hat{X}}(N/2^2), \ldots, \Phi_{\hat{X}}(N/2^h)$, B has already determined the bits $s_0, s_1, \ldots, s_{h-1}$. For the coefficient $\hat{X}(N/2^{h+1})$ one finds that

$$\Phi_{\hat{X}}(N/2^{h+1}) \overset{2\pi}{=} -\pi \frac{s}{2^h} + \Phi_{\hat{Y}}(N/2^{h+1})$$

$$\overset{2\pi}{=} -\pi s_h - \frac{\pi}{2^h}\{s\}_{2^h} + \Phi_{\hat{Y}}(N/2^{h+1}).$$

Now, clearly $\{s\}_{2^h} = s_{h-1} \cdots s_1 s_0$ is known to B, so that the only unknown term for B in the right hand side of the above equation is s_h. So, again, the sign of $\Phi_{\hat{X}}(N/2^{h+1})$ allows B to uniquely determine s_h. This proves that the $\log(S)$ bits that represent the signs of the phases $\Phi_{\hat{X}}(N/2^i), i = 1, \ldots, \log(S)$, allow B to reconstruct the value of s.

2.2 The concrete 2-D case

This technique described above can be easily extended so as to deal with 2-dimensional signals, i.e. the case where X and Y are two images. The solution is based on the separability of the DFT; we briefly describe here the main idea, the reader may find more details in [6]. Supposing the vertical and horizontal shift values are known to be restricted respectively to the intervals $[0, 2^p)$ and $[0, 2^q)$, under ideal conditions (i.e., noiseless images and circular shift) a good choice for the code associated to the shift information of X with respect to Y is given by

$$\Phi_{\hat{X}}(N/2^1, 0), \Phi_{\hat{X}}(N/2^2, 0), \ldots, \Phi_{\hat{X}}(N/2^p, 0), \quad (6)$$

$$\Phi_{\hat{X}}(0, N/2^1), \Phi_{\hat{X}}(0, N/2^2), \ldots, \Phi_{\hat{X}}(0, N/2^q). \quad (7)$$

The above expression clarifies the fact that under ideal assumptions, the separability of the DFT allows to detect shift between images, in general along a diagonal direction, as the composition of vertical and horizontal shifts. Thus, only the coefficients of the DFT associated with pure vertical and horizontal frequencies are used in this case to solve the problem.

Under non-ideal assumptions, however, this approach is not suitable and must be enforced. In particular, due to the unavoidable boundary effects of the relative shift between images, and to the additive noise that is usually present in every practical situation, it is necessary to consider a higher number of bits so as to add robustness to the code. In particular rather than using only the coefficients associated to vertical and horizontal frequencies, as in eq.'s (6), (7), we also consider "diagonal" frequency phases of the form $\Phi_{\hat{X}}(N/2^i, N/2^j)$. There is an interesting interpretation of this idea. By doing the proposed operation, in fact, we can consider the use of vertical and horizontal frequency coefficients as the encoding of the shift information; the encoding of the diagonal frequencies, then, can be considered as the enforcement of the extracted code by means of the evaluation of some "parity bits". For the decoding operation, different approaches have been considered in [6]. In particular, decoding techniques with different complexity, and thus with different efficiency, have been proposed and motivated. As a general

rule it is found in fact that for a given fixed amount of used bits, there is a trade off between the complexity of the decoder and the robustness of the code to the additive noise on the images.

3. ROTATION AND SCALE DETECTION USING THE FOURIER-MELLIN TRANSFORM

3.1 From shift to rotation and scale

The ideas recalled in the previous section for the distributed coding of relative shifts between images can be further investigated in the direction of an extension for the more general problem where two images do not only differ for a relative shift, but also for rotations and/or a scale factor. By properly operating on the DFT of the images, it is possible to reduce scale and rotations to a shift problem in a proper transformed domain. In the field of image registration this idea has been initially proposed in [11] and [12] for the problem of combined translations and rotations. The extension to the case of scale between images has then been studied in [13] and in [14] with the use of what is actually known as a Fourier-Mellin transform. We refer to [7] for a survey of image registration techniques and to [9, 10] for more details on the Fourier-Mellin transform and its applications in the field of signal processing.

Consider the case where the transformation between images X and Y is a combination of translation, rotation and scale. Consider for simplicity the case of noiseless images, where images are considered now as continuous domain signals. Thus, for n, m real numbers we can write the relation between the two images as

$$X(m,n) = Y(\lambda (m\cos\theta_0 + n\sin\theta_0) - r,$$
$$\lambda(-m\sin\theta_0 + n\cos\theta_0) - c) \quad (8)$$

By computing the Fourier transform of these two signals in the continuous frequency domain (i.e., for real k, l) we have

$$\hat{X}(k,l) = \frac{e^{-j\phi_{r,c}(k,l)}}{\lambda^2}\hat{Y}(\lambda^{-1}(k\cos\theta_0 + l\sin\theta_0),$$
$$\lambda^{-1}(-k\sin\theta_0 + l\cos\theta_0)) \quad (9)$$

where $\phi_{r,c}(k,l)$ is the phase term due to the translation by the vector $\mathbf{v} = (r,c)$. Now, by changing to log-polar coordinates, let us set $\bar{X}(\rho,\theta) = |\hat{X}(e^\rho\cos\theta, e^\rho\sin\theta)|$ and similarly $\bar{Y}(\rho,\theta) = |\hat{Y}(e^\rho\cos\theta, e^\rho\sin\theta)|$. Then, with some simple algebraic manipulations we have

$$\bar{X}(\rho,\theta) = \frac{1}{\lambda^2}\bar{Y}(\rho - \log\lambda, \theta - \theta_0) \quad (10)$$

So, the rotation and scale between the two images X and Y are reduced to a shift between the two signals $\bar{X}$ and $\bar{Y}$. Note that all the operations we have applied to the two images X and Y can be performed separately on each one; thus the operation on the image X can be performed by user A without any need to know Y. So, we have reduced the problem of communicating the rotation and scale parameters to the problem of a communicating shift in a proper domain and, for this problem, we can use the same technique described in the previous section. Once user B has recovered

the rotation and the scale factor, by rotating and rescaling the image Y it can obtain an image Y' that only differ from X by a shift, apart from the noise inevitably due to resampling operations. As last step, then user B can recover also the shift if it has been encoded with the distributed shift coding technique as described. The theoretical procedure derived in this section is analyzed from a practical point of view in the next subsection. The discussion developed, in fact, only holds rigorously in continuous space and frequency domains and, moreover, under the hypothesis of noiseless images defined on an infinite domain. For discrete images with limited support there are some issues to address in order to implement an algorithm that might work. In the next subsection a detailed analysis of such practical issues is given.

3.2 From the ideal case to the concrete problem

In this section a more practical description of the proposed technique is provided. As said, there are some differences between the theoretical setting studied in the previous sections and the concrete situation of real images. Here we give an overview of these differences and we comment on possible solutions to the problems. Figure 2 shows the global scheme for the proposed technique. A first important consideration for the use of the proposed ideas in a practical case is that the spectrum of the images X and Y, being the images are cropped versions from a same scene, is distorted by the effects of the window. For the case of natural images, the most relevant effect is usually that some false vertical and horizontal frequencies appear in the spectrum, which are due to the discontinuity that appear in the periodic replication of the image. In this case, if the two images content has a relative rotation, in the sense that the images are cropped with a relative rotation from an infinitely wide scene, then those false vertical and horizontal frequencies do not rotate with the remaining part of the spectrum, but instead stay in the vertical and horizontal direction. So, detecting rotations in the spectrum domain is difficult unless these frequencies are removed. In order to do that it is necessary to use a smoothing window to the two images X and Y (such as for example a Tukey window) so that their periodic replication do not contain significant false vertical and horizontal discontinuities. This operation was already suggested in [6], for the encoding of shifts, in order to remove the boundary effects. Thus it can remain unaltered, and it is even useful, for the detection of the rotational component.

Another practical problem is found in the change from Cartesian coordinates to Log-Polar coordinates. In this case, we must consider that a change in coordinates must be performed by opportune resampling in a practical scheme. The resolution used in this resampling step is very important in order to preserve the information about rotation and scale. In particular, the resolution used for the angular coordinate must be sufficient to detect the angular rotation with the required resolution. This is not really a problem, but care must be taken. A more important problem is the resampling required for the logarithmic reshaping of the radial coordinate. Due to the fact that we can only use integer coordinates in a digital representation of the image, the values of ρ in eq. (10) are in practice always integers; given that our algorithm for distributed coding of shifts is studied for the detection of integer shifts, we find that using that algorithm we can only estimate $\log\lambda$ to the nearest integer value. If the logarithms are taken
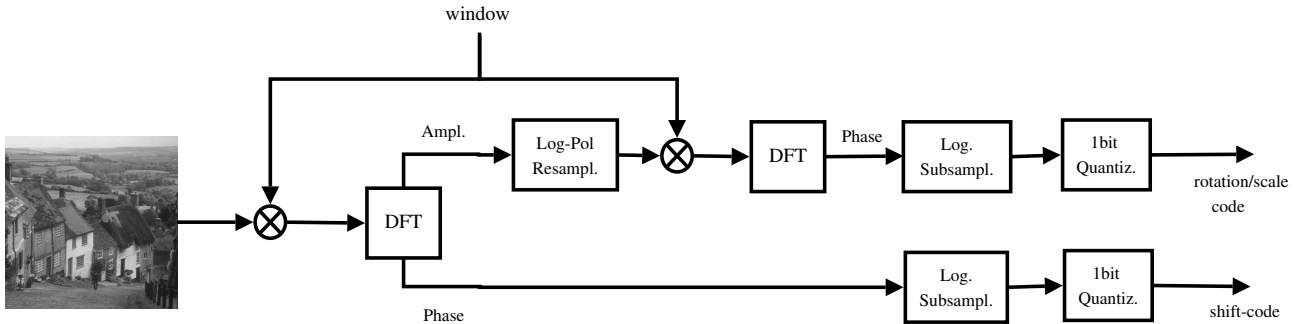
Figure 2: Scheme of encoding of the shift, rotation and scale parameters.

to the base e, the resolution for the scale factor is defined by the interval $[e^{\lfloor \log \lambda \rfloor}, e^{\lceil \log \lambda \rceil}]$ or, in other words, the scale factors that can be detected by the algorithm are the values $\ldots, e^{-2}, e^{-1}, 1, e, e^2, \ldots$. It is important to consider that typically interesting values for λ are values close to 1, and thus, using the base e cannot provide a good resolution for λ. In order to achieve such a good resolution it is necessary to use a small base μ for the logarithmic resampling, in particular a base μ sufficiently small so as to have two consecutive powers of μ sufficiently tight around λ. As a first approximation, we can consider that if we take a base $\mu = 1 + \varepsilon$, where ε is much smaller than 1, the detectable scale factors are the values $(1 + \varepsilon)^k$ with integer k, which can be approximated at the first order as $1 + k\varepsilon$. So, in order to have a resolution ε on the detectable scale factor we have to use a base $1 + \varepsilon$ in the logarithmic scale of the radial coordinate.

The last point that is important to clarify is that once the spectrum of the images are represented in the Log-Polar domain, in order to apply the distributed shift coding algorithm it is again necessary to apply a windowing in order to reduce the boundary effects. In this case, it is particularly important to perform this operation because the spectrum of natural images in the Log-Polar domain is mostly concentrated around the axis $\rho = 0$, and rapidly vanishes for high values of ρ. This causes strong boundary effects because of the fact that the ideal shifts for the use of the DFT phase are circular shifts, while here we have a non-circular shift. This is the same problem already encountered in [6], where the windows were applied directly on the images; here it is worth clarifying that the boundary effects are much more critical in this case, because of the particular shape of the Log-Polar Spectrum of natural images.

In the next section we give an example of how the proposed approach operates in order to recover the rotation and scale of a couple of images. In [6] a detailed study of the robustness of the method in the case of noisy images was proposed. For the problem of distributed coding of relative rotation and scale, with respect to the shift only problem, we found more difficult to properly test the proposed approach with noisy images. Even for noiseless images, which means images created by applying a rotation and a scale factor to a single initial image,[1] the choice of some parameters such as the resolution in the resampling operations, the type of

window applied to the image and to the spectrum in the Log-Polar domain, seem to have much more impact on the results. On the other hand, this is no surprise as the transformations applied to the images lead to an accumulation of boundary effects and sampling artefacts that make the algorithm less robust. So, for this problem it is necessary to consider possible ways of increasing the robustness by using more bits for the representation of the phase informations (consider that by applying the algorithm as presented for the shift problem, encoding rotation and scale requires only about 100 bits for 512x512 images). This aspect remains object of on-going research.

4. EXPERIMENTAL SIMULATION

In this section we use a simulation example to give a step by step description of the operations involved in the distributed coding of shift, rotation and scale information.

Consider two images X and Y as shown in Figure 3. The Image Y is obtained by rotating X of 10 degrees in the clockwise direction, and applying a scale factor of 1.25. We used these two images as test images in order to give a practical example of how the proposed technique works. The complete scheme of the performed operations undergone by the X image is shown in Fig. 2.

The image is first multiplied by a window (we used a Tukey window here), it is transformed with a DFT, and the phase is used for the encoding of the shift information as described in [6]. The amplitude of the DFT, instead, is undergone a resampling to obtain a representation in a Log-Polar domain. We have sampled both the θ and ρ coordinates with 512 samples. This means that the interval between two angular samples has length $\theta_{\min} = 0.7031$. The integer multiples of $\theta_{\min}$ that are closest to the real values of the rotation factor are $14\theta_{\min} = 9.8438$ and $15\theta_{\min} = 10.5469$. For the radial coordinate, then, we have chosen $\mu = 1.0086$. The integer powers of μ that are closest to the true scale factor are $\mu^{25} = 1.2393$, $\mu^{26} = 1.25$ and $\mu^{27} = 1.2608$. We remark here that once the base μ is chosen, one still has a choice in how to sample the ρ coordinate; more precisely, the 512 samples can be taken at values $\rho = \mu^{k+i}$, $i = 0, \ldots, 511$, where k is any relative integer. In order to have an effective algorithm it is convenient to choose k so as to spread the 512 samples in appropriate positions in the ρ axis. Here, for example we have chosen $k = 130$, which spreads the samples from $\mu^{130} = 3.0521$ to $\mu^{641} = 245.14$, which is an appropriate range given that the original image is a 512×512 image.

---

[1]Note that in this case there is actually at least a small noise due to the fact that images are resampled on different grids. It is clear that a good interpolation technique must be used to resample the images during the rotation and scaling operations in order to reduce this error.

(a) Image *X*.   (b) Image *Y*.

Figure 3: *X* and *Y* images used in this example. Here the relative rotation is 10 degrees, and the relative scale is 1.25.

(a) Registered version of Image *Y*.   (b) Prediction error after registration of *Y*.

Figure 4: Registered Image *Y* and the residual image obtained when using this registered version of *Y* to predict *X*.

Once the Log-Polar representation has been obtained by an appropriate resampling process, the spectra of the *X* image is undergone a second DFT transform. The obtained phase is then used together with the algorithm used for the encoding of shift information, in order to encode the rotation and scale factors. In this case, we have used a total amount of 100 bits for the encoding of the relative rotation and scale. The decoder, on its side, uses the algorithms described in [6] in order to recover these factors from the received bits. In this experiment, the decoded value of the rotation is $\theta' = 15\theta_{\min} = 10.5469$, which is very satisfactory, even if it is not optimal in terms of distance from the real rotation factor, as $14\theta_{\min} = 9.8438$ would be a better estimate. The scale factor is in this case "correctly" recovered as $\lambda' = \mu^{26} = 1.25$.

Thus, the rotation and scale factors are extracted by user *B*, and the inverse operations can be applied to the *Y* image in order to register the rotation and scale. The shift component is then detected using the procedure proposed in the [6], using the phase code of the *X* image sent by user *A*. The obtained shift-compensated image *Y* is shown in Figure 4(a).

In Figure 4(b) the "prediction" error, obtained when the registered image *Y* is used to predict the *X* image, is shown. This is important in a DVC setting. IN that context, indeed, we can consider the prediction error as the actual innovation of the image *X* with respect to the registered *Y* image. So, if we consider the problem of communicating *X* from user *A* to user *B* as in a Slepian-Wolf setting of distributed coding, once the registration of *Y* has been performed by user *B* as explained in this section, the amount of information to be transmitted reduces from what is shown in Figure 4(a)

to what is shown in Figure 4(b), which is a much smaller amount of information.

5. CONCLUSIONS AND PERSPECTIVES

In this paper we have formulated a problem of image registration with minimal information exchange, and in particular we have proposed a technique for the identification of the relative shift, rotation and scale factors between two images, available at different locations, with the exchange of a very small amount of bits. The proposed problem is of importance for example in applications such as Distributed Coding, where a decoder has to perform compensation of the available side information in order to extract a good approximation of the signal available to the transmitter. It is however possible to consider the problem studied in this paper as an interesting research field in its own right. The technique proposed here is a first possible solution to the particular case where the differences between the images are only related to shift, scale and rotations. It is clear that more general types of deformation between the images could be considered, for which it should be worth studying the possibility of performing a registration with minimal informatin exchange. The general topic is the object of ongoing research.

REFERENCES

[1] D. Slepian and J.K. Wolf. Noiseless coding of correlated information sources. *IEEE Trans. Inform. Theory*, 19: 471–480, 1973.

[2] N. Gehrig and P. L. Dragotti. DIFFERENT-DIstributed and Fully Flexible image EncodeRs for camEra sensor NeTworks. In *Proc. of IEEE Int. Conf. on Image Proc.*, Genoa, 2005.

[3] S. Aaron, S. Rane and B. Girod. Wyner-Ziv video coding with hash-based motion-compensation at the receiver. In *Proc. IEEE Int. Conference on Image Proc.*, Singapore, 2004.

[4] A. Aaron and B. Girod. Wyner-Ziv video coding with low-encoder complexity. In *Proc. Picture Coding Symposium*, PCS-2004, San Francisco, CA, 2004.

[5] M. Tagliasacchi and S. Tubaro. Hash-based motion modeling in Wyner-Ziv video coding. In *Proc. of the IEEE Int. Conf. on Acoust., Speech, and Signal Proc.*, Honolulu, Hawaii, 2007.

[6] M. Dalai, R. Leonardi and P. L. Dragotti. Distributed coding of shifts using the DFT phase. In *Proc. of the IEEE Int. Conf. on Acoust., Speech, and Signal Proc.*, Honolulu, Hawaii, 2007.

[7] L. G. Brown. A survey of image registration techniques. *ACM Computing Surveys*, 24(4):325–376, 1992.

[8] J. Altmann. On the Digital Implementation of the Rotation-Invariant Fourier-Mellin Transform. *Journal of Information Processing and Cybernetics*, EIK28:13–36, 1987.

[9] Y. Sheng, C. Lejeune, and H. H. Arsenault. Frequency-domain Fourier–Mellin descriptors for invariant pattern recognition. *Opt. Eng.*, 28:494–500, 1988.

[10] S. Derrode and F. Ghorbel. Robust and Efficient Fourier-Mellin Transform Approximations for Gray-Level Image Reconstruction and Complete Invariant Description. *Computer Vision and Image Understanding*, 83(1):57–78, 2001.

[11] S. Alliney, C. Morandi. Digital Image Registration Using Projections. *IEEE Trans. Pattern Anal. Mach. Intell.*, 8(2): 222-233, 1986.

[12] E. De Castro, C. Morandi. Registration of rotated and translated images using finite Fourier transforms. *IEEE Trans. Pattern Anal. Mach. Intell.*, 9(5): 700–703, 1987.

[13] Q. Chen, M. Defrise and F. Deconinck. Symmetric phase-only matched filtering of Fourier-Mellin transform for image registration and recognition. *IEEE Trans. Pattern Anal. Mach. Intell.*, 16: 1156–1168, 1994.

[14] B. Srinivasa Reddy and B. N. Chatterji. An FFT-based technique for translation, rotation and scale-invariant image registration. *IEEE Trans. Image Processing*, 5: 1266-1271, 1996.