Copula based Divergence Measures and their use in Image Registration

T S Durrani and Xuexing Zeng

Centre of Excellence in Signal and Image Processing
EEE Department, Royal College, Strathclyde University Glasgow, UK
Phone: +441415482883, fax: +441415522487, email: t.durrani@eee.strath.ac.uk; x.zeng@eee.strath.ac.uk

ABSTRACT

This paper explores a new measure, based on the copula density functions, for image registration, especially for the multimodal image registration. The measure relies on determining the mutual information between images taken at different times from different viewpoints or by different sensors. The process aims to find the optimal spatial correspondence that offers maximal dependence between the grey levels of the images when they are correctly aligned. Misalignment results in a decrease in the measure. To this effect, this paper focuses on improving the estimation of mutual information. It is shown that copulas form an integral definition of mutual information, and lead to robust estimation tools. The paper includes new results on generalised divergence measures, including the Kullback-Liebler divergence, Kolomgorov. Tsallis , I_α, and Renyi measures amongst others. These are expressed in terms of copula density functions. Results are presented on the registration of two classes of images, using the Clayton Copula to estimate the divergence between the images, and their performance evaluated.

1.0 Introduction

Issues of image registration arise in a number of areas, where information needs to be extracted from a multiplicity of scenes. The requirement to identify identical or quasi-identical elements in several images (or scenes) forms an important feature in the analysis and abstraction of information from data. Examples of such requirements arise in area such as medical imaging, non-destructive evaluation of faults, or remote sensing, amongst others. For instance there is a need to assess patient response to medication from tomographic images over a longitudinal treatment; study of the progression of a defect or fault from ultrasonic images taken at different intervals; erosion of coastline or forestry or land usage from satellite images or SAR data; and many more. In all the above cases, there is as need for comparing images, while recognising static or common feature, hence the need for image registration.

The problem of image registration may be encapsulated in the statement that multimodal image registration is a process of determining the optimal spatial correspondence between images that are taken at different times from different viewpoints or by different sensors.

To effect such registration, conventional techniques have relied upon the use of measures such as correlation techniques or Fourier and related methods. However in the nineties there was a paradigm shift in multimodality image registration, where intensity based algorithms were developed that exploited mutual information **[1--4]**. These were based on the assumption that there is maximal dependence between the grey values of the images when they are correctly aligned, misalignment was shown to result in a decrease in the measure, thus techniques were developed for improving the estimation of mutual information, as a precursor to image registration.

The well known definition for mutual information between two random variables X and Y is given by

$$I(X,Y)=\iint p(x,y)\log\frac{p(x,y)}{q(x)q(y)}dxdy \qquad (1)$$

where *p(x,y)* is the joint probability density function of the variables *(X,Y)*, and *q(x), q(y)* are the marginal densities of *X* and *Y* respectively.

According to the definition of mutual information, a key requirement is to estimate the joint density function for the images. There are two approaches to this estimation problem that may be classified as parametric methods or non-parametric methods. In the parametric method, a given form for the joint density function such as bi-variate Gaussian, bi-variate Gamma distribution **[5]** is assumed, and

© EURASIP, 2009

related parameters are estimated from the given set of finite number of data points, and the mutual information *I(X,Y),* is evaluated from this. However, the image distributions are rarely Gaussian nor Gamma distributed, and even have the different types of marginal distribution.

For non-parametric approaches, well-known methods rely on histogram and K Nearest Neighbor and Kernel (KNN) density estimation **[6]**. The main drawback of histogram method is that it requires a large amount of data for reliable estimates, and the KNN produces estimates with very heavy tails; and can have discontinuities. The resulting density is not a true probability density since its integral over all the sample space may diverge.

Copulas **[7]** offer alternative robust parametric techniques for the modeling of the dependency structure between random variables and hence within the observed data. Since copulas can model any joint multidimensional probability density function only from marginal density functions, hence are candidates for the evaluation of mutual information from observed data.

This paper will briefly introduce Copula functions, and their properties; extend their applications to a range of Divergence measures that include *I(X,Y)* as a special case, and then illustrate the results on real data to prove that the proposed copula approach provides robust estimates of the joint density function and mutual information, and as such offers a new technique for image registration.

2.0 Background to copulas

Copulas represent a mathematical relationship between the joint distribution and the marginal distribution of random variables. For two random variables X and Y, their joint probability distribution $F_{XY}(x, y)$, according to Sklar **[7]**, is given by:

$$F_{XY}(x,y) = C(F_X(x), F_Y(y)) = C(u, v)$$

where $C(u,v)$ is the copula function, $u=F_X(x)$, $v=F_Y(y)$ are marginal probability distributions, $F_{XY}(x,y)$ is the joint distribution, so the copula function be defined by:

$$C(u,v) = F_{XY}(F_X^{-1}(u), F_Y^{-1}(v))$$

Furthermore, the copula density is given by**[8]**:

$$c(u,v) = \frac{\partial^2 C(u,v)}{\partial u \partial v} = \frac{\partial^2 C(F_X(x), F_Y(y))}{\partial u \partial v}$$

$$= \frac{\partial^2 F_{XY}(x,y)}{f_X(x) f_Y(y) \partial x \partial y} = \frac{f_{XY}(x,y)}{f_X(x) f_Y(y)}$$

where $c(u,v)$ is the copula density function, $f_{XY}(x,y)$ is the joint density function of x and y and $f_X(x)$, $f_Y(y)$ are the marginal density function respectively.

3.0 Mutual information, Divergence and generalised measures

Mutual information as given in Eq.(1) is a measure of the distance between the joint density and the product of the marginal densities of two variables, and is known as a particular case of the 'divergence measure' call the Kullback-Leibler Divergence **[9]**.

In recent years a number of information theoretic similarity measures have been identified **[4] [10] [11].** These are based on generalised divergence measures and are known to offer improved accuracy, robustness and speedy convergence for image registration **[12].**

Using the definition of copula density functions as given above, it may be shown that the Kullback Liebler Divergence may be transformed into **[13]**

$$D_{KL} I(X,Y) = \int_0^1 \int_0^1 c(u,v) \log(c(u,v)) du dv \quad (2)$$

It may be seen that this expression is a much simpler manifestation that Eq. (1), and is completely defined in terms of the copula density function *{c(u,v)}*. It will be shown later that it offers ease of computation.

Other measures that have attracted recent attention are the Renyi divergence, Tsallis divergence, I_α divergence and Bhattacharya and Kolmogorov distances. Details of these may be found in **[12]**. It is interesting to note that each of these measures lends itself to definitions in terms of copula density functions.

Indeed, as shown below, the copula functions offer a more compact definition for the Generalised divergences, and thus provide an easier alternative for the computation of the mutual information of Divergence measures.

3.1 Generalised Divergence Measures

There are two types of Generalised Divergences **[14]**:

a. Czisar Divergence, and

b. Renyi Divergence

a. For Czisar Divergence, the copulas function may be used to derive a simpler and more compact definition, as

$$D = \int_0^1\int_0^1 \psi(c(u,v)dudv \quad (3)$$

Thus when $\psi(x) = x\log x$, Eq.(3) yields the Kullback –Liebler divergence measure. Similarly, when $\psi(x) = \frac{x^\alpha - \alpha x + \alpha - 1}{(\alpha - 1)}$, we obtain the Tsallis Divergence, while when $\psi(x) = \frac{x^\alpha - \alpha x + \alpha - 1}{\alpha(\alpha - 1)}$, Eq.(3) yields the I_α Divergence.

b. For Renyi Divergence family, we have the generalized expression:

$$D = \log(\psi^{-1}(\int_0^1\int_0^1 c(u,v)\psi(c(u,v)dudv)) \quad (4)$$

For $\psi(x) = x^{r-1}$, Eq.(4) yields the Renyi divergence included in Table 1, while $r = 1/2$, give the Bhattacharya divergence measure. Note that using L'Hopital Rule, the Renyi divergence reduces to the Kullback Liebler divergence as Table 1 gives an equivalence for the different divergence measures in terms of their density functions and related copula density functions. In all cases the measures may be computed from the copula density function alone.

To test the properties of the Divergence measures, the Clayton Copula density function was used to compute the various divergence measures **[12]**.

The Clayton copula distribution is defined as:

$$C(u,v) = (u^{-\theta} + v^{-\theta} - 1)^{-\frac{1}{\theta}}$$

and the Clayton copula density function is given by

$$c(u,v) = (1+\theta)u^{-1-\theta}v^{-1-\theta}(-1+u^{-\theta}+v^{-\theta})^{-2-\frac{1}{\theta}} \quad (5)$$

where copula parameter $(0 < \theta < \infty)$. The Clayton copula parameter θ can be related to Kendall's tau rank correlation **[7]** as:

$$\tau_{X,Y} = 4\iint_{[0,1]^2} C(u,v)dC(u,v) - 1 = \frac{\theta}{\theta + 2}$$

The Kendall tau rank correlation coefficient is a non-parametric statistic that is used to measure the degree of correspondence between two rankings and for assessing the significance of this correspondence **[7].**

For the Clayton copula density function, relationship between mutual information and Kendall rank correlation (τ) for different divergence measures are given in Figure 1. Note the convex relationship and the differences in performance between the various divergence measures.

Name	Divergence	By Copulas
Kullback - Leibler	$\iint p(x,y)\log(\frac{p(x,y)}{q(x)q(y)}dxdy$	$\iint_{[0,1]^2} c(u,v)\log[c(u,v)]dudv$
Kolmogorov	$\iint \lvert p(x,y) - q(x)q(y) \rvert dxdy$	$\iint_{[0,1]^2} \lvert c(u,v) - 1 \rvert dudv$
Tsallis	$\frac{1}{\alpha-1}(\iint \frac{p(x,y)^\alpha}{q^{\alpha-1}(x)q^{\alpha-1}(y)}dxdy - 1)$	$\frac{1}{(\alpha-1)}(\iint_{[0,1]^2} c(u,v)^\alpha dudv - 1)$
I_α	$\frac{1}{\alpha(\alpha-1)}(\iint \frac{p(x,y)^\alpha}{q^{\alpha-1}(x)q^{\alpha-1}(y)}dxdy - 1)$	$\frac{1}{\alpha(\alpha-1)}(\iint_{[0,1]^2} c(u,v)^\alpha dudv - 1)$
Renyi	$\frac{1}{(r-1)}(\log\iint \frac{p(x,y)^r}{q^{r-1}(x)q^{r-1}(y)}dxdy)$	$\frac{1}{(r-1)}\log\iint_{[0,1]^2} c(u,v)^r dudv$
Bhattcharyaa	$-2\log\iint \sqrt{p(x,y)q(x)q(y)}dxdy$	$-2\log\iint_{[0,1]^2} \sqrt{c(u,v)}dudv$

Table 1
Divergence by copulas

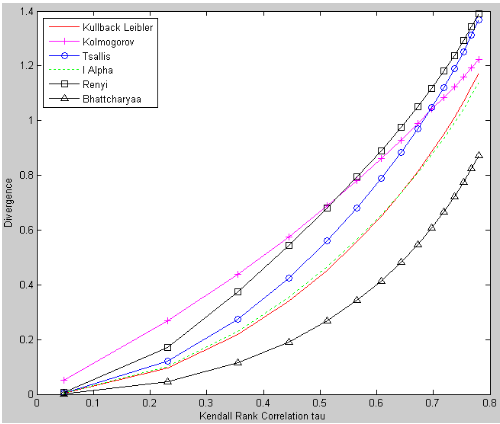

Figure 1(Mutual information vs Kendall rank correlation, where the parameter for Tsallis, I_α are 1.2 and for Renyi divergence is 1.5).

4.0 Image registration using copulas

To test the effectiveness of using copulas for image registration, a number of experiments were carried out on two difference classes of images – (i) Digital aerial photograph, and (ii) medical magnetic resonance images. Two images from the same class were employed - one the original image and the second a rotated version of the original with added 2% salt and pepper noise. The objective was to determine the alignment (registration) between the images by computing the different divergence measures, using the Clayton copula function. The alignment is achieved for the value of the rotated angle when the divergences reach a maximum.

(i) Image registration for aerial images.

Figure 2 (a)
Reference Image

Figure 2 (b)
Floating Image

Figure 2 (c)
Registered image

Figure 2 (d)
Figure 2 (c) overlaid on (a)

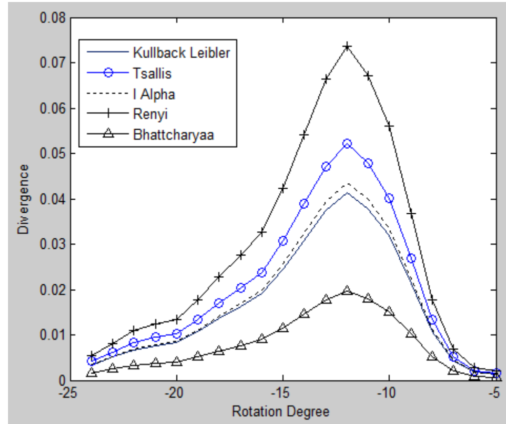

Figure 2 (e)

The aerial images in Figure 2 (a) and 2 (b) are available from the Matlab website **[15]**, taken at different times and from different viewpoints. Using the Clayton copula model, the copula parameter θ is estimated for the two images, drawing upon the method of Canonical Maximum Likelihood (CML) as proposed in **[16]**. Fig 2(c) is the registered image computed from Figure 2(b). Figure 2(d) is generated by overlaying a semitransparent version of the registered image Figure 2(c) on the Figure 2(a). The Divergence Measures given in Table 1 were calculated with *{c(u,v)}* as given in Eq (3), and were computed for the Tsallis, I_α (parameter α =1.2) and Renyi families (r=1.5). The computation values of the divergence for different rotational angles are given in Figure 2 (e). It may be seen that all the Divergence values peaks when the rotation angle is 12.1^0 clockwise – (the correct answer).

(ii) Image registration for magnetic resonance images.

The Magnetic resonance images were obtained from **[17]** and are used to determine image registration using the Clayton copulas. Figure 3(a) is the original image, while Figure 3(b) is a rotated version of Figure 3(a) with added 2% “salt and pepper” noise. Figure 3(c) is the registered image from (b). Figure 3(d) is generated in a similar measure to Figure 2(d) by overlaying the registered image Figure 2(c) on Figure 2(a).

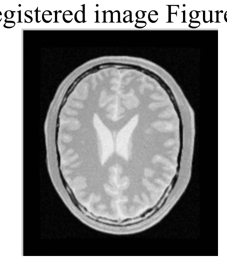

Figure 3 (a)
Reference image

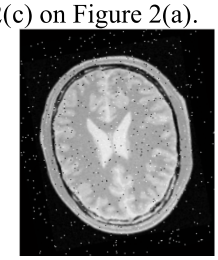

Figure 3 (b)
Floating image

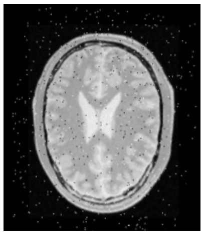

Figure 3 (c)
Registered image

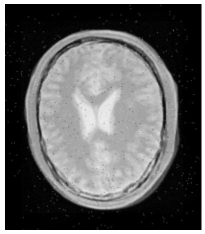

Figure 3 (d)
Figure (c) overlaid on (a)

The Divergences were computed by using Clayton copula. Here the parameter α =1.2 was used for Tsallis, I_α divergences, while Renyi divergence for *r=1.5* was computed. As may be seen from Figure 3(e), the Divergences reach a single maximum when the rotation angle is 9.25^0 clockwise, and all the Divergence measures are very similar.

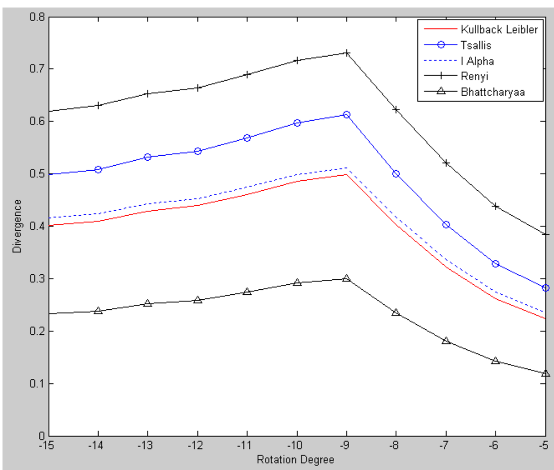

Figure 3 (e)

5.0 Conclusions

In this paper image registration techniques based on copula functions has been presented. Copulas have been used to calculate a number of divergences for image registration. The mutual information computed, and used to compare several divergence measures. These are then evaluated using the Clayton copula and applied to two sets of images to illustrate their efficacy in image registration. It is shown that the copula measure offers a more robust way to register images.

6.0 References

[1] F. Maes, A. Collignon, D.Vandermeulen, G.Marchal and P. Suetens, "Multimodality image registration by maximization of mutual information", IEEE Trans, Med. Image, vol 16, no.2, pp 187-198, Apr, 1997.

[2] W.M. Wells, P. Viola, H. Atsumi, S. Nakajima and R. Kikinis, "Multi-modal volume registration by maximization of mutual information", Med. Image Anal, vol 1, no. 1, pp. 35-52, Mar. 1996.

[3] J.P.W. Plluim, J.B.A. Maintz, and M. A. Viergever, "Mutual information based registration of medical images: A survey", IEEE Trans Med. Image, vol 22, no.8 pp. 986-1004, Aug. 2003.

[4] J.P.W. Plluim, J.B.A. Maintz, and M. A. Viergever, "f information measures in medical image registration", IEEE Trans, Med. Image, vol.23, no. 12, pp. 1508-1516, Dec, 2004

[5] F.Chatelain, J.Y Tourneret, J.Inglada, A.Ferrare, "Bivariate Gamma distribution for image registration and change detection", IEEE Trans, Image Processing, Vol 16, issue 7, pp. 1796-1806, July 2007.

[6] K. Fukunaga, "Introduction to statistical pattern recognition", Elservier academic press, 1999.

[7] B.Roger, "An Introduction to copulas", Springer series in statistics, 1999.

[8] T.S Durrani and X. Zeng, "Copulas for bivariate probability distribution", Electronics Letters, IET, Vol 43, Issue:4 pp.248-249,2007.

[9] S. Kullback, "The Kullback-Leibler distance", The American Statistician vol 41,pp. 340-341.

[10]Y.He, A.B. Hamza and H.Krim, "A generalized divergence measure for robust image registration", IEEE Trans. Signal Process, vol 51, no. 5, pp. 1211-1220, May 2003.

[11] S. Martin, G. Morison, W. Nailon and T. Durrani, "Fast and accurate image registration using Tsallis entropy and simultaneous perturbation stochastic approximation", Electron. Lett, vol 40, no. 10, pp. 595-597, May 2004.

[12] S. Martin, T.S. Durrani, "A new divergence measure for medical image registration", IEEE Trans, image processing, Apr 2007.

[13] M. Davy, A. Doucet, "A new insight into positive Time-Frequency distribution", IEEE signal processing letters, 2003.

[14] I. Csiszar, "Information-type measures of difference of probability distributions and indirect observations", Studia Sci. Math. Hungarica, vol. 2, pp. 299-318, 1967.

[15]http://www.mathworks.com/access/helpdesk/help/toolbox/images

[16] V. Durrleman, A. Nikeghbali and T. Roncalli, "Which copula is the right one?", Groupe de Recherche Operationnelle, Credit Lyonnais, working paper, 2000.

[17]http://www.doc.ic.ac.uk/~dr/software/index.html