SELF-ASSESSING IMAGE-BASED RESPIRATORY MOTION COMPENSATION FOR FLUOROSCOPIC CORONARY ROADMAPPING

Michael Manhart[⋆†] *Ying Zhu*[⋆] *Dime Vitanovski*[†]

[⋆]Siemens Corporate Research, 755 College Road East, Princeton, NJ, USA
[†]Pattern Recognition Lab, Friedrich-Alexander University, Erlangen, Germany

ABSTRACT

We present a self-assessing image-based motion compensation method for coronary roadmapping in fluoroscopic images. Extending our previous work on respiratory motion compensation, we introduce kernel-based nonparametric data analysis in this work to better characterize the objective function involved in motion estimation, which leads to two new improvements in motion compensation. First, through mode analysis we are able to capture the dominant component of the respiratory image motion and increase the chance of finding the global optimum. Second, an information theoretic measure is proposed to assess the uncertainty of the motion estimation and automatically detect unreliable motion estimates. The benefits of the proposed method are shown through evaluations performed on real clinical data from different procedures of percutaneous coronary interventions.

Index Terms— Image Motion Compensation, Uncertainty Analysis, Fluoroscopic Angiography, Image Guidance, Coronary Roadmapping

1. INTRODUCTION

As the primary imaging modality for cardiac intervention, live X-ray fluoroscopy is widely used in percutaneous coronary intervention (PCI) procedures, for instance, for directing guidewires, placing stents and crossing chronic total occlusions (CTO). Dynamic coronary roadmapping [1] and multimodality image fusion [2] are important techniques developed to provide advanced imaging guidance, where 2D or 3D vessel roadmaps acquired from fluoroscopic, CT or MR angiography are superimposed on live fluoroscopy for real-time guidance. In both cases, reliable motion compensation is critical for dynamically moving 2D or 3D coronary roadmaps to properly match live fluoroscopy.

To deal with cardiac motion, ECG gating is commonly used to synchronize live images with corresponding roadmaps at approximately the same cardiac phases. In contrast, respiratory motion is less reproducible and drawbacks are associated with respiratory gating [3]. The effect of respiration on the heart and coronary arteries as well as representing motion models have been investigated in earlier studies involving volumetric and biplane data [3, 4, 5]. It remains a challenging issue to compensate breathing motion in live fluoroscopy, which consists of transparent layers of static bone tissues and moving soft tissues as well as contrast-filled vessels and devices. In a previous work [6], Zhu et al. introduced a method to identify static structures from fluoro images over a cardiac cycle, which allowed the image information of soft tissues to be separated and used for respiratory motion estimation. However, the gradient-based approach adopted by [6] remains susceptible to the local optimum problem when the matching function exhibits multiple local optima. In this paper, we extend the work of [6] and make two new contributions. First, through sampling of the motion parameter space and introducing nonparametric data analysis techniques, we are able to identify the dominant component of the respiratory motion to overcome the local optimum problem. Second, we propose a new entropy-based measure to assess the uncertainty of motion estimation and automatically detect unreliable motion estimates.

2. METHODS

2.1. Problem Formulation

Different motion models including translation, rigid body and affine transformation have been studied in earlier works [4, 3, 5] to characterize the respiratory motion effect on the heart. Denote a reference image as $I_R(\mathbf{x})$ and an incoming image with approximately the same cardiac phase acquired at time t as $I_t(\mathbf{x})$. The problem is to estimate the soft tissue motion $W(\mathbf{x}, P)$ of the heart under breathing between I_R and I_t with parameters P. This is formulated as an optimization problem in our previous work [6],

$$P = \operatorname{argmax} \sum_{\mathbf{x}} \kappa_t(\mathbf{x}) \kappa_R(W(\mathbf{x}; P)) f(I_t(\mathbf{x}), I_R(W(\mathbf{x}, P))) \tag{1}$$

where κ_R and $\kappa_t : \mathbb{R}^2 \rightarrow [0,\ 1]$ are weighting functions introduced to exclude the areas with contrast wash-in and wash-out between I_R and I_t as well as to suppress static image structures in motion estimation. Details on how to calculate κ_R and κ_t are provided in [6]. The weights deal with static structures such as spines, surgical devices, etc.. Ribs are moving

structures during breathing. In cases where their projected image motion is negligible, the proposed method can still be applied. Nevertheless, how to handle multiple moving structures remains an open issue in general.

Sum of squared difference (SSD) is used in [6] to define the matching function f. To be more robust to changes of image contrast caused by varying acquisition conditions, the weighted normalized cross correlation (WNCC) is adopted here as the objective function.

$$P = \text{argmax}\, f_{wncc}(P) \qquad (2)$$
$$f_{wncc}(P) = \frac{cov_w(I_t(\mathbf{x}), I_R(W(\mathbf{x};P)))}{\sqrt{var_w(I_t(\mathbf{x}))var_w(I_R(W(\mathbf{x};P)))}}$$

where $cov_w(I_t(\mathbf{x}), I_R(W(\mathbf{x};P))) = E_w[I_t(\mathbf{x})I_R(W(\mathbf{x};P))] - E_w[I_t(\mathbf{x})]E_w[I_R(W(\mathbf{x};P))]$ is the weighted covariance between $I_t(\mathbf{x})$ and $I_R(W(\mathbf{x};P))$, $var_w(I_t(\mathbf{x})) = E_w[I_t(\mathbf{x})^2] - (E_w[I_t(\mathbf{x})])^2$ and $var_w(I_R(W(\mathbf{x};P))) = E_w[I_R(W(\mathbf{x};P))^2] - (E_w[I_R(W(\mathbf{x};P))])^2$ are the weighted variances of $I_t(\mathbf{x})$ and $I_R(W(\mathbf{x};P))$ respectively with $E_w[I_t(\mathbf{x})I_R(W(\mathbf{x};P))] = \frac{1}{\mathcal{K}}\sum_{\mathbf{x}} \kappa_t(\mathbf{x})\kappa_R(W(\mathbf{x};P))I_t(\mathbf{x})I_R(W(\mathbf{x};P))$, $E_w[I_t(\mathbf{x})] = \frac{\sum_{\mathbf{x}} \kappa_t(\mathbf{x})I_t(\mathbf{x})}{\sum_{\mathbf{x}} \kappa_t(\mathbf{x})}$, $E_w[I_R(W(\mathbf{x};P))] = \frac{\sum_{\mathbf{x}} \kappa_R(W(\mathbf{x};P))I_R(W(\mathbf{x};P))}{\sum_{\mathbf{x}} \kappa_R(W(\mathbf{x};P))}$, and $\mathcal{K} = \sum_{\mathbf{x}} \kappa(\mathbf{x})\kappa(W(\mathbf{x};P))$.

2.2. Global Optimization through Kernel-based Nonparametric Data Analysis

The gradient descend algorithm used in [6] is susceptible to the local optimum problem. When the objective function has multiple local optima, the algorithm is not guaranteed to converge to the global optimum. To overcome this limitation, we apply the nonparametric data analysis technique [7] to characterize the objective function f_{wncc} in the parameter space and identify the global optimum among multiple local optima.

Nonparametric data analysis is widely used in data mining and computer vision to effectively analyze complex data distributions with multiple modes. To apply this technique to optimization, we first perform sparse sampling in the parameter space and obtain an approximation of the objective function f_{wncc} using kernel approximation. Assume that $\{p_1, \cdots, p_n\}$ are a set of sparse samples in the parameter space with function values $\{f_{wncc}(p_1), \cdots, f_{wncc}(p_n)\}$. An approximation of f_{wncc} is obtained using the kernel density approximation technique developed in [8].

$$\tilde{f}_{wncc}(P) = \sum_{i=1}^{n} w_i K(P; p_i, \sigma_s^2) \quad (w_i \geq 0) \qquad (3)$$

A common choice for $K(\cdot)$ is the Gaussian kernel $K(P; p_i, \sigma_s^2) = \frac{\exp(-||P-p_i||^2/(2\sigma_s^2))}{(2\pi\sigma_s^2)^{dim(P)/2}}$ defined by its center p_i and bandwidth σ_s. The non-negative weights w_i are determined by minimizing the sum of squared differences between $\{\tilde{f}_{wncc}(p_i)\}$ and $\{f_{wncc}(p_i)\}$. This leads to a non-negative least squares problem for which an efficient solver is provided by [9]. Essentially, $\tilde{f}_{wncc}$ is a smoothed version of f_{wncc} [7]. The well-known mean shift algorithm [10] is applied to find all the modes, i.e. local maxima, of $\tilde{f}_{wncc}$, which are denoted as $\{m_1, \cdots, m_M\}$. In the second step, dense sampling is performed in the local neighborhood $N(m_i)$ around each mode m_i and function values $\{f_{wncc}(p_{i,1}), \cdots, f_{wncc}(p_{i,n_i})\}$ are computed and normalized over the dense sample set $\{p_{i,1}, \cdots, p_{i,n_i}\}$,

$$\bar{f}(p_{i,j}) = \max\{0, f_{wncc}(p_{i,j}) - f_{m,i} + 0.1\} \qquad (4)$$

where $f_{m,i} = \max\{f_{wncc}(p_{i,j}) : j = 1, \cdots, n_i\}$. For each mode m_i, an improved local approximation $\hat{f}_i(P)|_{P\in N(m_i)}$ is obtained from the normalized function values $\{\bar{f}(p_{i,j})\}$ at the dense samples $\{p_{i,j}\}$ through kernel approximation.

$$\hat{f}_i(P) = \sum_{j=1}^{n_i} \hat{w}_j K(P; p_{i,j}, \sigma_d^2) \quad (\hat{w}_j \geq 0) \qquad (5)$$

The mean shift algorithm is used to locate the mode $\hat{m}_i$ of the improved local approximation $\hat{f}_i$. Finally, the mode with the highest function value is returned as the global optimum, $m_{opt} = m_{i^*}$, where $i^* = \text{argmax}\, f_{wncc}(\hat{m}_i)$. The normalization of $\{\bar{f}(p_{i,j})\}$ is necessary to make the confidence measure proposed in the next subsection comparable. In addition, by suppressing the function value for P away from the local mode, the kernel approximation is able to focus on a close neighborhood of the local mode.

In our experiments, global optimization was performed in the translational space, i.e. $P = [t_x, t_y]'$, as translational motion is the dominant component of breathing motion observed in coronary fluoroscopy. However, the proposed global optimization algorithm applies to other parametric motion models as well.

2.3. Uncertainty Analysis and Self Assessment

By exploring the information about the shape of the matching function, we can further assess the uncertainty associated with the motion estimation. Surfaces of f_{wncc} with a single dominant mode as shown in Fig. 1(a) indicate less ambiguity in the image data and more reliable matches. In contrast, flat surfaces or surfaces with multiple modes as shown in Fig. 1(b) indicate large ambiguity in the image data and large uncertainty associated with motion estimates.

To quantify the uncertainty of motion estimation, we use a similar technique as presented in [11] to fit a Gaussian distribution to the surface of the matching function $\hat{f}_i(P)$ in the local neighborhood $N(\hat{m}_i)$ of a mode $\hat{m}_i$.

$$\hat{f}(P)|_{P\in N(\hat{m}_i)} \approx K(P; \hat{m}_i, \hat{C}_i) \qquad (6)$$
$$= \frac{\exp(-\frac{1}{2}(P-\hat{m}_i)^T \hat{C}_i^{-1}(P-\hat{m}_i))}{(2\pi)^{dim(P)/2}|\hat{C}_i|^{1/2}}$$

We calculate the covariance matrix $\hat{C}_i$ by minimizing the Euclidean distance between the surface gradients of the improved approximation $\hat{f}$ and the Gaussian kernel K in local neighborhood $N(\hat{m}_i)$

$$\hat{C}_i = \underset{C \in \mathcal{S}}{\operatorname{argmin}} \sum_{j=1}^{n_i} ||\nabla K(P; p_{i,j}, C) - \nabla \hat{f}_i(p_{i,j})||^2 \quad (7)$$

where $\mathcal{S}$ denotes the set of positive definite matrices. This least squares problem over positive definite matrices is solved by the method presented in [12]. Examples for the local Gaussian fitting are shown in Fig. 2. A well-known measure for quantifying the uncertainty of random variables is entropy. The differential entropy of a Gaussian distributed random variable with covariance matrix C is defined as [13]

$$h(C) = \ln(2\pi e) + 0.5 \cdot \ln(|C|) \quad (8)$$

We omit the constant terms and define the uncertainty of the matching function around mode $\hat{m}_i$ as

$$\epsilon_i = \ln(|\hat{C}_i|) \quad (9)$$

To determine the uncertainty of the complete surface, we combine the uncertainty measures for multiple modes through a weighted sum. Modes with higher function values f_{wncc} are weighted higher and minor modes with lower function values are ignored. The entropy based uncertainty measure is defined as

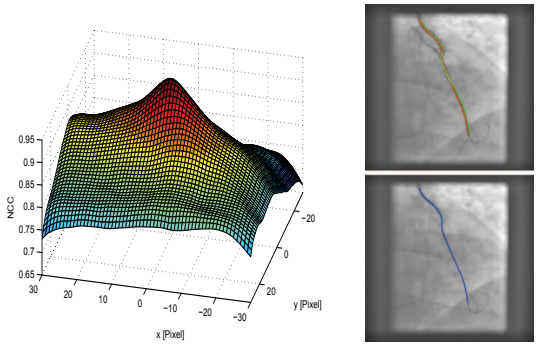
$$\begin{aligned} \Upsilon &= \textstyle\sum_{i=1}^{M} \omega_i \epsilon_i \quad (10) \\ \omega_i &= \max\{0, 1 - (f_{wncc}(m_i) - f_m)/s_{th}\} \\ (f_m &= \max\{f_{wncc}(\hat{m}_i) : i = 1, \cdots, M\}) \end{aligned}$$

where s_{th} is a threshold. We use Υ for the self-assessment of the optimization algorithm. A motion estimate is accepted when its uncertainty measure is below a threshold Υ_{th}, and the matching function value is above a threshold f_{th}. Otherwise, the motion estimate is rejected.

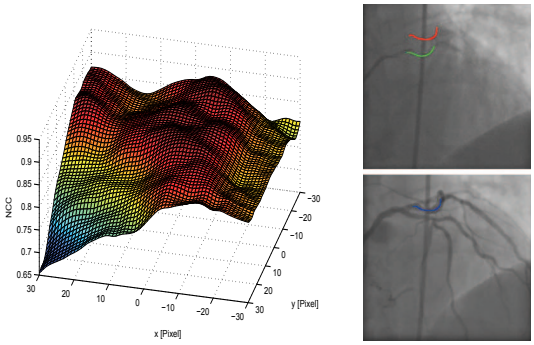
3. RESULTS

We used fluoroscopic data from 16 CTO cases to evaluate the global optimization algorithm and the self-assessment technique. The data was acquired by Angiographic C-arm systems (AXIOM Artis, Siemens Healthcare), with pixel size ranging from 0.184 mm to 0.216 mm. These cases were chosen because they had either vessel structures visible or guidewires present which provided the ground truth of vessel centerline for evaluation. Visible vessel structures as well as guidewires in the fluoroscopy data were annotated as splines to represent vessel centerlines.

As Fig. 1 shows, for every test case, multiple images associated with the same cardiac phase were annotated. One of



(a) Function with single dominant mode ($\Upsilon = 6.85$).

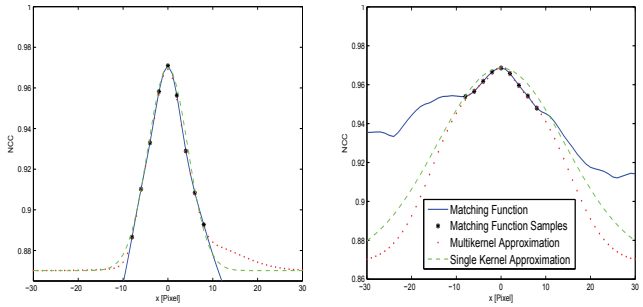


(b) Function with multiple modes ($\Upsilon = 51.28$).

Fig. 1. Surfaces of different matching functions.

the images was selected as the reference image (lower images in Fig. 1), and its annotated spline (shown in blue) simulated the initial roadmap overlay. The rest of the images were used for testing (upper images in Fig. 1) with the corresponding vessel structures annotated as ground truth (shown in green) for evaluation. We performed motion estimation between the reference image and each testing image, and used the estimated motion to transform the annotated spline in the reference image to match the test image. The transformed spline simulated the motion compensated overlay (shown in red). To quantify the motion estimation error, the distance between the motion compensated overlays (red splines) and the ground truth structures (green splines) was calculated. We calculated point to spline distances between each pair of splines, and the median of all distances was chosen as the error distance.

Global optimization was performed on fluoroscopy images at a reduced resolution. Sparse sampling in the translational space was performed in the range of ± 32 pixels with a sampling distance $\Delta t_x = \Delta t_y = 8$ pixels for the smooth approximation $\tilde{f}_{wncc}$, and in the range of ± 8 pixels with a distance $\Delta t_x = \Delta t_y = 2$ pixels for the improved approximations $\{\hat{f}_i\}$. The bandwidth of the Gaussian kernels was set to $\sigma_s^2 = 64$ for $\tilde{f}_{wncc}$ and $\sigma_d^2 = 4$ for $\{\hat{f}_i\}$. $\Upsilon_{th} = 14$, $f_{th} = 0.8$ and $s_{th} = 0.05$ were used to perform self assessment on motion estimates.



(a) Dominant mode $\epsilon = 6.39$. (b) Flat mode $\epsilon = 10.59$.

Fig. 2. Fitting a Gaussian distribution to the matching function around a mode.

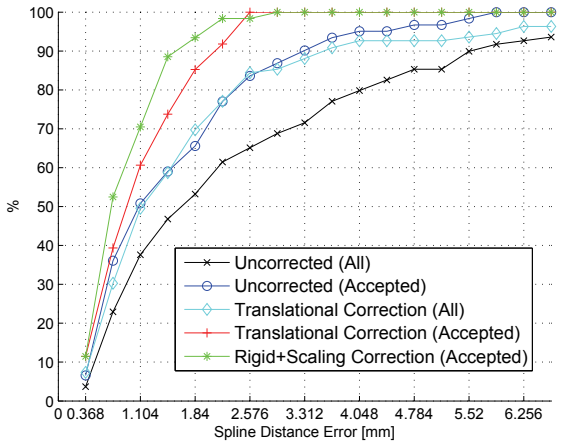


Fig. 3. Cumulative distribution of motion compensation error.

#	Accepted Frames	Uncorrected (All)	Uncorrected (Acc.)	Translational (All)	Translational (Acc.)	R+S (Acc.)
1	0/5 (0 %)	3.407	N/A	1.567	N/A	N/A
2	9/10 (90 %)	0.572	0.503	1.299	1.365	0.564
3	5/6 (83 %)	0.869	0.695	0.800	0.612	0.554
4	4/4 (100 %)	0.818	0.818	1.212	1.212	1.026
5	0/7 (0 %)	2.386	N/A	2.177	N/A	N/A
6	2/8 (25 %)	1.406	0.470	1.547	1.130	0.630
7	5/5 (100 %)	1.984	1.984	1.254	1.254	1.340
8	5/5 (100 %)	1.420	1.420	1.133	1.133	0.913
9	9/9 (100 %)	1.935	1.935	1.455	1.455	1.226
10	5/10 (50 %)	2.128	1.368	0.959	0.693	0.874
11	3/9 (33 %)	4.905	3.121	3.702	0.763	0.811
12	1/6 (17 %)	3.360	1.882	1.965	2.050	1.580
13	1/7 (14 %)	1.261	1.196	3.345	2.408	2.788
14	2/7 (28 %)	6.592	1.164	3.579	1.012	0.883
15	7/7 (100 %)	2.328	2.328	0.990	0.990	1.023
16	3/4 (75 %)	4.791	3.618	0.635	0.570	0.744
Tot.	61/109 (59 %)	2.474	1.564	1.796	1.121	0.948

Mean Error Distances in mm.

To further improve the results we used a rigid+scaling (R+S) motion model, extending the translational model by a rotation and two scaling parameters in x and y direction. To find the optimal parameters we applied a gradient descent based optimization strategy as discussed in [6] initialized with accepted results of the global optimization. Table (1) shows, for each of the 16 test sequences, the number of test matches, the number of accepted matches and the mean of the error distance from the globally optimized translational model, the R+S model and without any motion correction. Without motion correction, the mean error over all motion estimates is 2.474 mm. With a translational motion model, the mean error reduces to 1.796 mm without self-assessment, and to 1.121 mm with self-assessment. With a R+S motion model and self-assessment, the error reduces to 0.948 mm. By performing self assessment, 59% of the motion estimates are accepted.

Figure 3 shows a cumulative distribution describing the percentage of tests with an estimation error equal or below the values on the x axis under different motion models. Without motion compensation, $\sim$ 53% of the motion estimates have an error of 1.84 mm or below. With translational motion compensation, $\sim$ 70% of all motion estimates and $\sim$ 85% of the accepted motion estimates have an error of 1.84 mm or below. Under the R+S motion model, this percentage increases to 93%. The average computation time for finding the translational parameters using the global optimization algorithm was 198 ms. The algorithm was implemented in C++ using OpenMP. The testing machine had a dual core Intel Pentium 4 processor with 2 $\times$ 2.8 GHz, 3 GB RAM and Windows XP SP3 as operating system.

4. DISCUSSIONS

In this paper, we presented a new image-based motion compensation method with self assessment for coronary roadmapping in fluoroscopic images. We extended our previous work in soft tissue motion estimation and introduced the techniques of kernel-based function approximation and mode analysis to overcome the local optimum problem. Furthermore, an entropy-based uncertainty measure was developed to automatically assess the reliability of motion estimation. Reliable motion compensation is not always possible, for instance, when the diaphragm or lung tissues obscure the heart or the soft tissue of the heart is only weakly visible. Our experience showed, that in these cases the shape of the matching function tended to be flat or have multiple modes. Exploring the information about the shape of the matching function helps to identify unreliable motion estimates. The evaluation of the global optimization algorithm and the self-assessment method showed, that all the cases with incorrect motion estimation were detected and almost all accepted motion estimates had an error below 2.6 mm compared to the annotated ground truth motion. However, the acceptance rate in some cases is very low. Future study includes combining the self-assessment technique with local motion models to explore motion in different subregions, which may provide a way for robust motion estimation in these difficult cases. In addition, the global optimization and self assessment techniques dis-

cussed in this work can be extended to more complex motion models.

5. REFERENCES

[1] JL Elion, "Dynamic coronary roadmapping," *US Patent 4878115*, October 1989.

[2] D Ruijters, BM ter Haar Romeny, and P Suetens, "Vesselness-based 2d 3d registration of the coronary arteries," *Int J CARS*, vol. 4, no. 4, pp. 391–397, June 2009.

[3] K McLeish, DLG Hill, DB Atkinson, JM Blackall, and R Razavi, "A study of the motion and deformation of the heart due to respiration," *IEEE TMI*, vol. 21, no. 9, pp. 1142–1150, 2002.

[4] AP King, R Boubertakh, KS Rhode, YL Ma, P Chinchapatnam, G Gao, T. Tangcharoen, M Ginks, M Cooklin, JS Gill, DJ Hawkes, RS Razavi, and T. Schaeffter, "A subject-specific technique for respiratory motion correction in image-guided cardiac catheterization procedures," *Med. Image Anal.*, vol. 13, no. 3, pp. 419–431, 2009.

[5] G Schechter, Ozturk C, JR Resar, and ER McVeigh, "Respiratory motion of the heart from free breathing coronary angiograms," *IEEE TMI*, vol. 23, no. 8, pp. 1046–1056, 2004.

[6] Y Zhu, Y Tsin, H Sundar, and F Sauer, "Image-based respiratory motion compensation for fluoroscopic coronary roadmapping," *MICCAI 2010 Part III*, pp. 287–294, September 2010.

[7] BW Silverman, *Density estimation for statistics and data analysis*, Chapman and Hall, 1986.

[8] B Han, Y Zhu, D Comaniciu, and LS Davis, "Visual tracking by continuous propagation in sequential bayesian filtering framework," *IEEE TPAMI*, vol. 31, no. 5, pp. 919–930, May 2009.

[9] J Cantarella and M Piatek, "tsnnls: A solver for large sparse least squares problems with non-negative variables," 2004, http://arxiv.org/abs/cs/0408029v1.

[10] Y Cheng, "Mean shift, mode seeking, and clustering," *IEEE TPAMI*, vol. 17, no. 8, pp. 790–799, August 1995.

[11] Kevin Nickels and Seth Hutchinson, "Estimating uncertainty in ssd-based feature tracking," *Image and Vision Computing*, vol. 20, pp. 47–58, 2002.

[12] Y Chen and JE McInroy, "Estimation of symmetric positive definite matrices from imperfect measurements," *IEEE TAC*, vol. 47, no. 10, pp. 1721–1725, October 2002.

[13] NA Ahmed and DV Gokhale, "Entropy expressions and their estimators for multivariate distributions," *IEEE TIT*, vol. 35, no. 3, pp. 688–692, May 1989.