2D Projective Transformations.

Manuel Felipe Pineda Loaiza

***Abstract*—In this work I applied the theory about 2D projective transformations studied in the computer vision class. The work consists of three experiments, the first experiment aims to remove the perspective of an image, the second one simulates a scanner, and the last one applies the perspective to one image in order to embed it into another image.**

***Keywords*—Computer vision, 2d projective transformations, 2d geometry, homographies, perspective.**

I. Introduction

A projective transformation is a natural deformation of the geometrical shapes that we see every day, for example when we take photografies of doors, windows and similar objects, they do not look like their original shapes (rectangles) but our brain automatically removes the perspective and we can know how the objects really look. This is a very interesting problem that we can solve using linear algebra and computational geometry, and it is in fact part of the basics for more complex tasks in the computer vision.

In order to define a projective transformation, we need to define first the homogeneous coordinates:

In mathematics, homogeneous coordinates or projective coordinates, introduced by August Ferdinand Mbius in his 1827 work Der barycentrische Calcl, are a system of coordinates used in projective geometry, as Cartesian coordinates are used in Euclidean geometry. They have the advantage that the coordinates of points, including points at infinity, can be represented using finite coordinates. [1]

Formally a projectivity is is an invertible mapping h from IP 2 to itself such that three points x 1 , x 2 and x 3 lie on the same line if and only if h(x 1), h(x 2) and h(x 3) do.

A planar projective transformation is a linear transformation on homogeneous 3-vectors represented by a non-singular 3 x 3 matrix

A. Homography

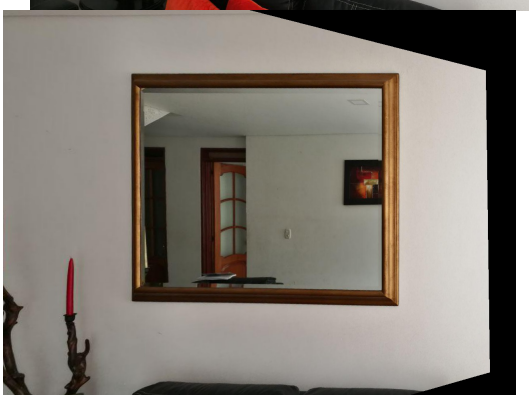
[2]

B. Bilinear interpolation

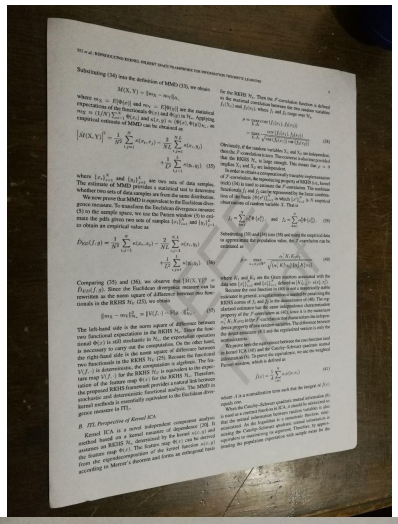
[3]

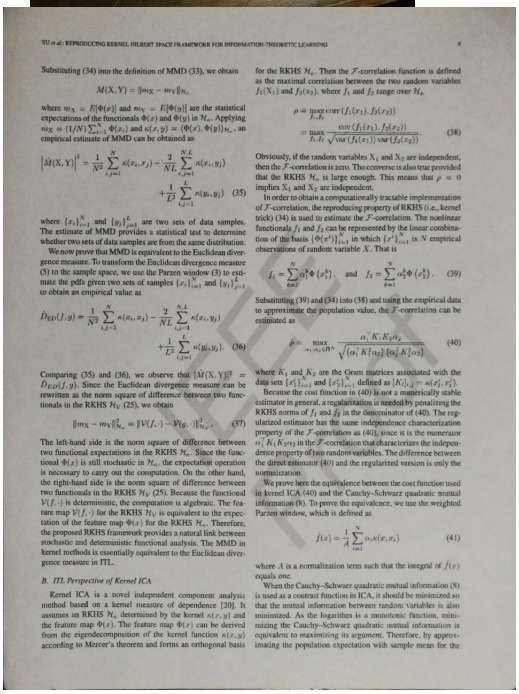
II. Experiments

A. Remove perspective



B. Scanner





Substituting (34) into the definition of MMD (33), we obtain

$$M(X,Y) = \|m_X - m_Y\|_{\mathcal{H}},$$

where $m_X = E[\Phi(x)]$ and $m_Y = E[\Phi(y)]$ are the statistical expectations of the functionals $\Phi(x)$ and $\Phi(y)$ in $\mathcal{H}_\kappa$. Applying $m_X = (1/N)\sum_{i=1}^{N}\Phi(x_i)$ and $\kappa(x,y) = \langle\Phi(x),\Phi(y)\rangle_{\mathcal{H}_\kappa}$, an empirical estimate of MMD can be obtained as

$$\left[M(X,Y)\right]^2 = \frac{1}{N^2}\sum_{i,j=1}^{N}\kappa(x_i,x_j) - \frac{2}{NL}\sum_{i,j=1}^{N,L}\kappa(x_i,y_j) + \frac{1}{L^2}\sum_{i,j=1}^{L}\kappa(y_i,y_j) \quad (35)$$

where $\{x_i\}_{i=1}^{N}$ and $\{y_i\}_{i=1}^{L}$ are two sets of data samples. The estimate of MMD provides a statistical test to determine whether two sets of data samples are from the same distribution.

We now prove that MMD is equivalent to the Euclidean divergence measure. To transform the Euclidean divergence measure (5) to the sample space, we use the Parzen window (3) to estimate the pdfs given two sets of samples $\{x_i\}_{i=1}^{N}$ and $\{y_i\}_{i=1}^{L}$ to obtain an empirical value as

$$D_{ED}(f,g) = \frac{1}{N^2}\sum_{i,j=1}^{N}\kappa(x_i,x_j) - \frac{2}{NL}\sum_{i,j=1}^{N,L}\kappa(x_i,y_j) + \frac{1}{L^2}\sum_{i,j=1}^{L}\kappa(y_i,y_j). \quad (36)$$

Comparing (35) and (36), we observe that $\left[M(X,Y)\right]^2 = D_{ED}(f,g)$. Since the Euclidean divergence measure can be rewritten as the norm square of difference between two functionals in the RKHS $\mathcal{H}_\Omega$ (25), we obtain

$$\|m_X - m_Y\|_{\mathcal{H}}^2 = \|V(f,\cdot) - V(g,\cdot)\|_{\mathcal{H}_V}^2. \quad (37)$$

The left-hand side is the norm square of difference between two functional expectations in the RKHS $\mathcal{H}_\kappa$. Since the functional $\Phi(x)$ is still stochastic in $\mathcal{H}_\kappa$, the expectation operation is necessary to carry out the computation. On the other hand, the right-hand side is the norm square of difference between two functionals in the RKHS $\mathcal{H}_\Omega$ (25). Because the functional $V(f,\cdot)$ is deterministic, the computation is algebraic. The feature map $V(f,\cdot)$ for the RKHS $\mathcal{H}_\Omega$ is equivalent to the expectation of the feature map $\Phi(x)$ for the RKHS $\mathcal{H}_\kappa$. Therefore, the proposed RKHS framework provides a natural link between stochastic and deterministic functional analysis. The MMD in kernel methods is essentially equivalent to the Euclidean divergence measure in ITL.

B. ITL Perspective of Kernel ICA

Kernel ICA is a novel independent component analysis method based on a kernel measure of dependence [20]. It assumes an RKHS $\mathcal{H}_\kappa$ determined by the kernel $\kappa(x,y)$ and the feature map $\Phi(x)$. The feature map $\Phi(x)$ can be derived from the eigendecomposition of the kernel function $\kappa(x,y)$ according to Mercer's theorem and forms an orthogonal basis for the RKHS $\mathcal{H}_\kappa$. Then the $\mathcal{F}$-correlation function is defined as the maximal correlation between the two random variables $f_1(X_1)$ and $f_2(x_2)$, where f_1 and f_2 range over $\mathcal{H}_\kappa$

$$\rho = \max_{f_1,f_2}\mathrm{corr}\left(f_1(x_1), f_2(x_2)\right) = \max_{f_1,f_2}\frac{\mathrm{cov}\left(f_1(x_1), f_2(x_2)\right)}{\sqrt{\mathrm{var}\left(f_1(x_1)\right)\mathrm{var}\left(f_2(x_2)\right)}}. \quad (38)$$

Obviously, if the random variables X_1 and X_2 are independent, then the $\mathcal{F}$-correlation is zero. The converse is also true provided that the RKHS $\mathcal{H}_\kappa$ is large enough. This means that $\rho = 0$ implies X_1 and X_2 are independent.

In order to obtain a computationally tractable implementation of $\mathcal{F}$-correlation, the reproducing property of RKHS (i.e., kernel trick) (34) is used to estimate the $\mathcal{F}$-correlation. The nonlinear functionals f_1 and f_2 can be represented by the linear combination of the basis $\{\Phi(x^i)\}_{i=1}^{N}$ in which $\{x^i\}_{i=1}^{N}$ is N empirical observations of random variable X. That is

$$f_1 = \sum_{i=1}^{N}\alpha_1^i\Phi\left(x_1^i\right), \quad \text{and} \quad f_2 = \sum_{i=1}^{N}\alpha_2^i\Phi\left(x_2^i\right). \quad (39)$$

Substituting (39) and (34) into (38) and using the empirical data to approximate the population value, the $\mathcal{F}$-correlation can be estimated as

$$\hat{\rho} = \max_{\alpha_1,\alpha_2\in\mathbb{R}^N}\frac{\alpha_1^\top K_1 K_2\alpha_2}{\sqrt{\left(\alpha_1^\top K_1^2\alpha_1\right)\left(\alpha_2^\top K_2^2\alpha_2\right)}} \quad (40)$$

where K_1 and K_2 are the Gram matrices associated with the data sets $\{x_1^i\}_{i=1}^{N}$ and $\{x_2^i\}_{i=1}^{N}$ defined as $[K_i]_{j,l} = \kappa(x_i^j, x_i^l)$.

Because the cost function in (40) is not a numerically stable estimator in general, a regularization is needed by penalizing the RKHS norms of f_1 and f_2 in the denominator of (40). The regularized estimator has the same independence characterization property of the $\mathcal{F}$-correlation as (40), since it is the numerator $\alpha_1^\top K_1 K_2\alpha_2$ in the $\mathcal{F}$-correlation that characterizes the independence property of two random variables. The difference between the direct estimator (40) and the regularized version is only the normalization.

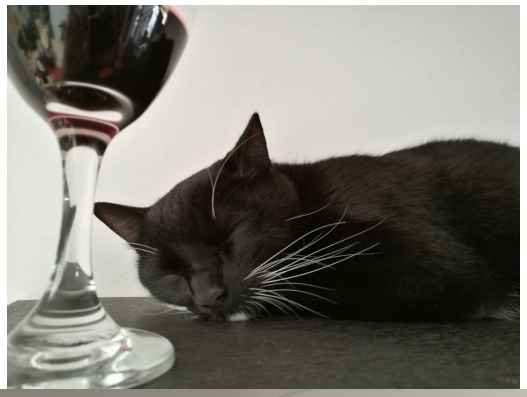
We prove here the equivalence between the cost function used in kernel ICA (40) and the Cauchy–Schwarz quadratic mutual information (8). To prove the equivalence, we use the weighted Parzen window, which is defined as

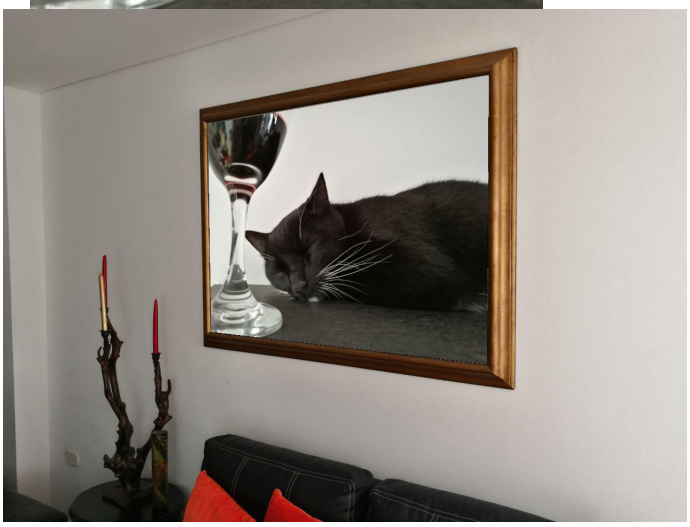
$$\hat{f}(x) = \frac{1}{A}\sum_{i=1}^{N}\alpha_i\kappa(x,x_i) \quad (41)$$

where A is a normalization term such that the integral of $\hat{f}(x)$ equals one.

When the Cauchy–Schwarz quadratic mutual information (8) is used as a contrast function in ICA, it should be minimized so that the mutual information between random variables is also minimized. As the logarithm is a monotonic function, minimizing the Cauchy–Schwarz quadratic mutual information is equivalent to maximizing its argument. Therefore, by approximating the population expectation with sample mean for the

C. Add perspective





III. Conclusion

It was cool :D

References

[1] Wikipedia. Homogeneous coordinates — wikipedia, the free encyclopedia, 2017. [Online; accessed 17-September-2017].

[2] Richard Hartley and Andrew Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, New York, NY, USA, 2 edition, 2003.

[3] Wikipedia. Bicubic interpolation — wikipedia, the free encyclopedia, 2017. [Online; accessed 16-September-2017].