# RECOVERING PLANAR PROJECTIONS OF PRINTED CLUSTERED-DOT HALFTONES

*Stephen Pollard, Robert Ulichney and Matthew Gaubatz*

**Hewlett Packard Laboratories.**

## ABSTRACT

This paper presents a novel method to recover planar projections (homographies) from perspectively distorted images of clustered-dot halftones. The method uses the variation of the local affine deformation of the halftone (estimated using a Fourier method) to recover the additional projective components. Evidence provided from simulated distortion and mobile capture shows that the approach provides an effective first step in the problem of using mobile device to recover stenographic information embedded in unknown halftone images, as pixel-accurate alignment is needed to recover the data.

*Index Terms*— Fourier Transform, Planar Projection, Homography, Vanishing Points, Halftone, Stegatone.

## 1. INTRODUCTION

We are interested in recovering undistorted images of printed clustered-dot halftones using a mobile device such as a mobile phone with a high resolution camera module. As the quality and resolution of mobile phone modules continues to improve, cameras are able to more clearly delineate the individual dots of a printed halftone pattern. In particular our focus is on using mobile devices to recover stenographic information that has been represented in the halftone structure by modulating the position of the individual printed halftone dots [1]. In order to achieve this goal it is necessary to correct the inevitable perspective distortion of the image caused by the non-frontoplanar orientation of the camera.

Previously [2] we have relied on having knowledge of the image used to create the data-bearing halftone and have used a multi-scale gradient descent alignment scheme [3] derived from the well-established Lucas and Kanade method [4] to recover affine and projective distortions; an initial approximation can be achieved by matching image features similar to SIFT [5] recovered from the coarsest scale of the multi-scale representation. There is an advantage for some applications, however, if knowledge of the image used to generate a data-bearing halftone remains unknown to the recovery device. Such a recovery scheme requires that the planar projective distortion be recovered directly from the image of the printed halftone, a problem solved in the work presented here.

If we are able to accurately distinguish the four corners of a rectangular printed image of known aspect ratio, then it is straight forward to compute the planar homography that relates the captured image to the original print by solving a linear system of equations with 8 unknowns [6]. In practice we may not know the aspect ratio of the image, its corners may not be preserved in the halftone rendering process (as they may lie in a highlight region of the image and hence halftone dots may not extend to the corners) or the extent of the physical print may extend beyond the limits of the captured image. An alternative is to use a scheme akin to that used to dewarp documents [7, 8] where text grouping is typically used to form first words and then lines, and to group bundles of lines and/or identify orthogonal edges such as justified paragraph boundaries, and from these identify two vanishing points that define the projective plane. This process, however, is made more complex for halftones by the fact that the local structure within the halftone pattern is more difficult to discriminate as it also depicts the pictorial content and as a result it can be computationally expensive to use low-level grouping strategies to effectively recover the vanishing points directly from the halftone patterns.

Instead, to our knowledge for the first time, we use the periodic nature of the halftone pattern and its representation in Fourier domain to robustly compute affine approximations to the local transform at different points in the captured halftone image, and from their subtle combination recover the remaining parameters of the homography up to a translation. Our method assumes that the structure and size of the halftone screen is known in advance, which is not unreasonable as we are in control of the printing process. If, however, the size is not known, but the structure is (e.g. that it is a standard 45 degree halftone screen), then there will be an additional unknown global scale parameter in the recovered planar perspective transformation.

## 2. METHOD

In order to remove the projective distortion from the perspective image of a plane it is necessary to recover the projective transform or homography that relates the rectified image coordinates to their distorted counterparts. In homogeneous coordinates this transform can be represented as a 3x3 matrix, $H$

$$\begin{pmatrix} x' \\ y' \\ w' \end{pmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{pmatrix} x \\ y \\ w \end{pmatrix} = \begin{bmatrix} A & t \\ v^T & 1 \end{bmatrix} \begin{pmatrix} x \\ y \\ w \end{pmatrix} \qquad (1)$$

or simply $\mathrm{x}' = H\mathrm{x}$, where finally $x'' = x'/w'$ and $y'' = y'/w'$. Note that the planar homography is the most general form of this transformation and that it can be represented as being composed from affine $A$ (2x2) and translation $t$ components when the elements of $v$ are zero.

For simplicity we restrict our attention to a monochrome 45 degree clustered-dot halftone screen with a cell period of 150 cells per inch (when printed at 600 dots per inch; dpi) in each of the horizontal and vertical directions (referred to as a 106.1 line per inch (lpi) screen as the spacing is measured in the 45 degree screen

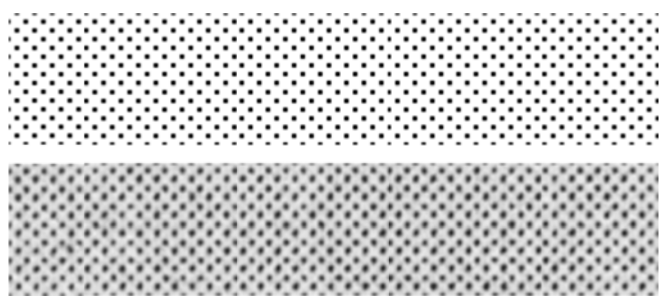

Figure 1. Digital 45 degree 12.5% halftone screen (top) and a similar scanned print from a HP LaserJet 1200 (bottom).

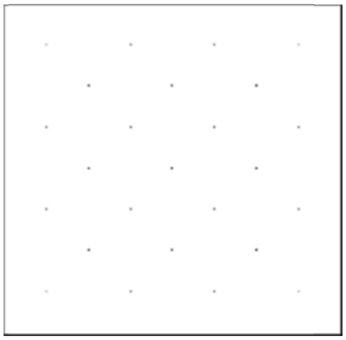

Figure 2. Magnitude of the Fourier transform (shown as inverted log to see weaker higher frequency components) of the 45 degree halftone screen in Figure 1.

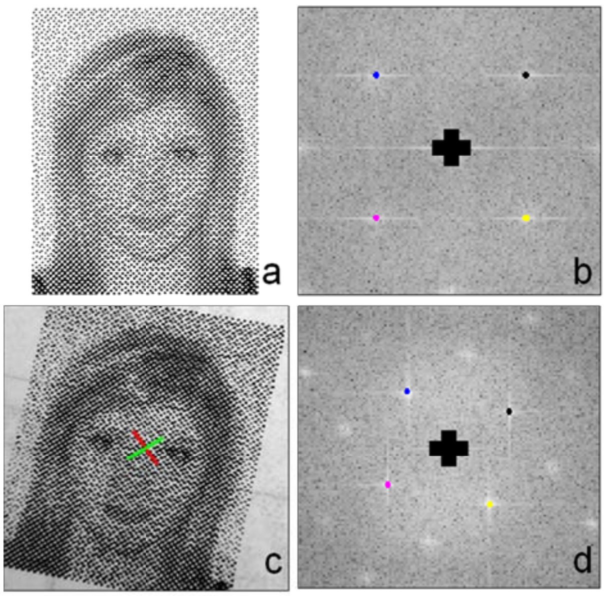


Figure 3. (a) Digital bitmap of a Stegatone with (b) its Fourier transform highlighting fundamentals using color dots; (c) a region of a 400dpi print with its Fourier Transform in (d).

direction). An example of the regular structure of this screen is shown in Figure 1 for a 12.5% uniform grey level (e.g. value 224 out of a range 0 to 255). This screen is the most commonly used monochrome clustered-dot halftoning screen (for example here shown for the print from a HP LaserJet 1200) and is the same structure as we use in our work on data-bearing halftones, or Stegatones [1]. Computing the discrete Fourier Transform (DFT) of a patch of this 12.5% grey halftone pattern and examining its magnitude (as shown in Figure 2) we see that the fundamental frequencies of the halftone pattern along the 45 degree directions are strongly visible as a set of 4 peaks surrounding the DC (at the center of the DFT). Also evident are a number harmonics repeated toward the boundary of the baseband.

### 2.1 AffineDFT

We find that this structure of the DFT is strongly preserved when the halftone depicts image content, even in the case where the structure of the halftone has been disturbed to create a Stegatone. Figure 3 shows a Stegatone of a face image and the magnitude of the Fourier transform of a central 256x256 image region in both its pure 316x400 pixel digital form and after printing (at 400 dpi) and mobile capture using a 2MP Web Camera (where the depicted region is approximately 600x600 pixels square). Each cell in the halftone is 4x4 pixels with many of the cells/dots shifted to encode information. It is straight-forward to recover the affine components of the transform (except for the translation which must be set to zero) from the relative displacement of the fundamentals using the well-known result [9] that an affine transform measured in the DFT is related to that in the image space as

$$A_I = (A_{FT}^{-1})^T \qquad (2)$$

where $A_I$ and $A_{FT}$ are the 2x2 affine transforms in the image and Fourier domains respectively. $A_{FT}$ can be solved by identifying the corresponding locations of a pair of fundamentals between the digital and captured images (e.g. 3(b) and 3(d)). Note that the Fourier Transform is symmetric with respect to the location of the DC and thus appropriate pairs of fundamentals describe all four locations. We call this the AffineDFT method. What is more, in those cases where the actual transform includes planar perspective components, AffineDFT will provide a reasonable approximation

$$\mathrm{x}' = \begin{pmatrix} x' \\ y' \\ w' \end{pmatrix} = \begin{bmatrix} A_I & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} x \\ y \\ w \end{pmatrix} = H_A \mathrm{x} \qquad (3)$$

to the local planar homography at the center of the region covered by the DFT. The accuracy of this approximation will depend on the magnitude of the non-affine components and as the transform becomes more and more strongly perspective (shorter viewing distance and less frontoparallel) then the approximation will break down as the locations of the fundamentals, in the DFT, becomes less distinct.

To help illustrate the local affine approximation to the transform, we show overlaid in Figure 3(c) diagonal red and green lines that correspond to the 45 degree screen directions (with respect to the original digital halftone pattern). As we move about the captured image the transformed screen directions vary to reflect the non-affine nature of the planar projection.

### 2.3 ProjDFT

Consider now the case where we know the full planer homography $H$ that transforms points from the digital halftone pattern into the captured image. This transformation is illustrated in Figure 4 where we consider the North, East, West and South cardinal points ($N, E, W, S$) displaced a distance $M$ from an origin $O$. The diagonals that join the cardinal points correspond to the 45 degree screen directions of the halftone. In this case the positions of the cardinal points in the perspective image are simply

$$N' = HN; \; E' = HE; \; W' = HW; \; S' = HS \qquad (4)$$

as shown on the right in Figure 4; the projected diagonals can be extended to meet at vanishing points $VP1$ and $VP2$ on the line at infinity. Lines parallel to one screen direction will converge on $VP1$ and lines parallel to the other converge on $VP2$.

Conversely, given the vanishing points $VP1$ and $VP2$ and selecting any two points as approximations for $N'$ and $S'$, the intersecting lines from the vanishing points through them will produce $E'$ and $W'$ where the quadrilateral $N'E'S'W'$ is guaranteed to be the projection of a rectangle in the original halftone space the sides of which are parallel to the screen directions. The closer the approximation of $N'$ and $S'$ to their true locations, the closer the rectangle will be to the original square defined by cardinal points $NESW$ in the rectified halftone space. In any case, the planar homography that relates the four points $NEWS$ and $N'E'W'S'$ will be correct up to an unknown scale and aspect ratio (needed to map the unknown rectangle to the known square).

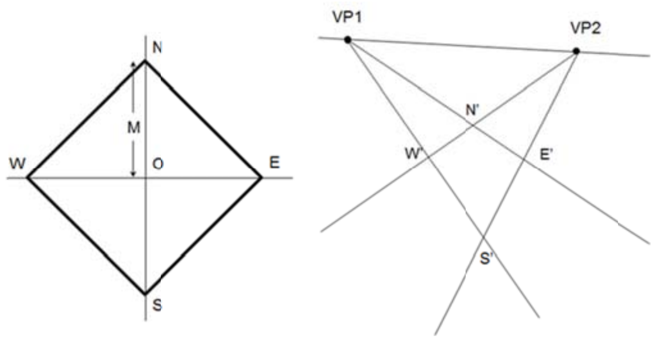


Figure 4.

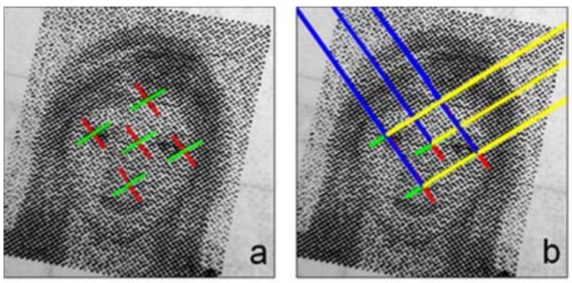


Figure 5. (a) Transformed cardinal points and recovered screen directions; (b) Vanishing lines for those same points.

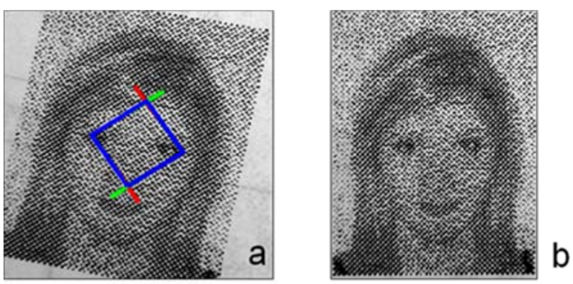


Figure 6. (a) Reconstructed East and West cardinal points using approximate North and South cardinal points plus calculated vanishing points; (b) Dewarped image using homography calculated from the 4 transformed cardinal points.

We solve the homography relating the $NEWS$ and $N'E'W'S'$ quadrilaterals using a standard linear method [6] as non-linear methods hold no advantage when computing an exact solution. That is, we form an 8x9 matrix $A$ where each matched pair of points $X$ and $X'$ contribute two rows

$$\begin{pmatrix} 0 & 0 & 0 & x_1x'_3 & x_2x'_3 & x_2x'_3 & x_1x'_2 & x_2x'_2 & x_3x'_2 \\ x_1x'_3 & x_2x'_3 & x_3x'_3 & 0 & 0 & 0 & -x_1x'_1 & -x_2x'_1 & -x_3x'_1 \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \\ h_5 \\ h_6 \\ h_7 \\ h_8 \\ h_9 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad (5)$$

where the length-9 vector $h$ is formed from the elements of $H_V$ in row order. The solution of the $h$ up to an unknown scale corresponds to the null space of $A$ and can be found using single value decomposition ($A=UDV^T$ where $D$ is a diagonal matrix of singular values) where $h$ then corresponds to the column of $V$ with the smallest single value in $D$. The final scale and aspect ratio can be recovered by applying the AffineDFT method to a reconstructed image based on $H_V$ to reveal a correcting affine transform $A_C$

$$H_C = \begin{bmatrix} A_C & 0 \\ 0 & 1 \end{bmatrix} H_V \quad (6)$$

The overall approach, called ProjDFT, is as follows:

1. Use AffineDFT to estimate an initial central affine $A_I$.
2. Use $A_I$ to transform the cardinal points to $N'E'W'S'$ (Equations 3 & 4 and Figure 5a).
3. At each transformed cardinal point, apply AffineDFT to recover the local screen direction vectors (Figure 5a).
4. Approximate two vanishing points $VP1$ and $VP2$ from the sets of orthogonal screen directions (Figure 5b).
5. Update $E'$ and $W'$ by intersecting lines from $VP1$ and $VP2$ through the original $N'$ and $S'$ (Figure 6a).
6. Calculate the homography $H_v$ from the new positions of the four cardinal points (Equation 5).
7. Rectify the original image using $H_v$ (Figure 6b).
8. Apply AffineDFT to recover $A_c$ and, in turn, the final corrected homography $H_c$ (Equation 6).

In practice, we can iterate ProjDFT to successively improve the approximation as the use of AffineDFT to approximate the local screen directions improves for the increasingly small perspective distortions. The vanishing points are calculated by minimizing the orthogonal error at the end of the appropriate 45 degree screen direction vectors for the line from the vanishing point through the respective transformed cardinal point.

## 3. RESULTS

We have conducted a large number of experiments on simulated data in order to explore the utility of the ProjDFT algorithm across a wide range of image types and system parameters. We have also performed a more limited test on real printed data comparing the performance to an existing image registration technique [2].

### 3.1 Simulated Data

In order to explore a wide parameter space we construct a large number of 2Kx2K digital halftones and related Stegatones derived from them with random payloads. We use 492 images from the McGill calibrated color image database [10] (specifically, the Animals, Landscapes and Man Made subsections). From each 786x576 RGB color TIFF images we extract the central 500x500 region of the green channel, scale it to 2000x2000 pixels and generate halftone and random payload Stegatone images with 4x4 pixels to each halftone cell.

Distorted test images are generated by selecting a halftone or Stegatone at random from the pool of 492 and warping it using a random planar homography of prescribed magnitude. The homography is defined by randomly displacing the four corners of the image over a range $[-R, +R]$ according to a uniform random distribution in each of the $x$ and $y$ directions. The homography is estimated using ProjDFT, with no knowledge of the pattern other than the fact that it is derived from a 45 degree halftone with a cell size of 4x4 pixels, and compared directly against the randomly generated homography used to warp the original halftone. The transforms themselves will not be identical as ProjDFT does not recover the translation. Thus, instead, we measure the extent to which they transform the relative locations of the image close to where the transform was estimated (i.e. the center of the distorted test image) back to the original rectified halftone geometry. In order to compare the two (inverse) homographies $H'$ and $H''$, the following relative error measurement is used. Consider two points $C$ and $R = C{+}D$ where $C$ is the center of the distorted halftone image and $R$ is a relative displacement $D$ from it. The respective transformed image locations are

$$C' = H'C;\ R' = H'R;\ C'' = H''C;\ R'' = H''R \quad (7)$$

from which the relative displacements after transformation are

$$D' = R' - C';\ \ D'' = R'' - C'' \quad (8)$$

and the relative error is $D''\text{-}D'$ and can be expressed as a relative percentage error

$$E_{\%} = \frac{100 * \|D'' - D'\|}{\min(\|D'\|, \|D''\|)} \quad (9)$$

For a tested image, the final reported accuracy of the homography is obtained by computing the value $E_{\%}$ at the four corners of a 200x200 pixel square centered on $C$, and taking the maximum.

Typical results for a DFT of size 512x512 and a spacing of cardinal points, $M$, of 512 pixels is shown in the graph in Figure 7 for 10 iterations of ProjDFT, plotting the median $E_{\%}$ (over 100 random trials) for distortion parameter, $R$, ranging from 40 to 400 pixels. Note that for all values of $R$, the median error reduces to about 0.13% after the 10 iterations asymptotically approaching a limit governed by the number of samples in the DFT and the resolution of the image. For comparison a similar experiment using the existing image based registration method has a median percentage error rate of 0.08%.

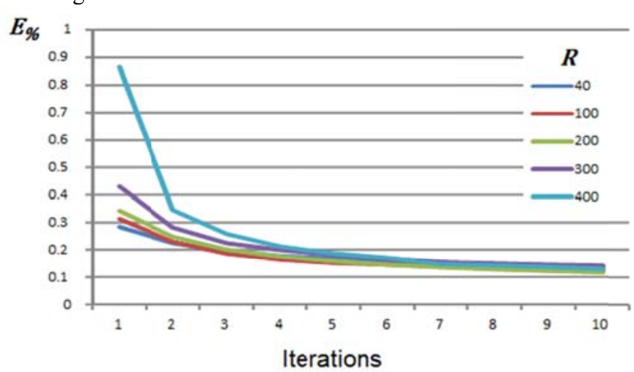


Figure 7. Relative error for various levels of random distortion ($R$) over 10 iterations of ProjDFT.

In Table 1, we explore the two most important parameters of the ProjDFT method: the size of the DFT and the spacing, $M$, for a fixed distortion $R$ = 400 pixels. For each parameter combination we present the median percentage error, $E_{\%}$, after 10 iterations. Notice, that as the size of the DFT and the spacing of the cardinal points increase from 128/64 to 512 pixels the accuracy improves significantly. Note also, that all results presented here are for Stegatone data; the results for un-modified halftones are similar.

Table 1.Median percentage error over 100 tests for various sizes of the DFT (rows) and spacing, $M$, of the cardinal points (columns).

| $E_{\%}$ | 64 | 128 | 256 | 512 |
|---|---|---|---|---|
| 128 | 0.96 | 0.68 | 0.53 | 0.59 |
| 256 | 0.47 | 0.37 | 0.26 | 0.25 |
| 512 | 0.29 | 0.25 | 0.19 | 0.13 |

### 3.1 Mobile Capture Data

The real data consisted of 100, randomly selected, Stegatone images printed at 400dpi on a monochrome HP LaserJet 4345M printer. Each was captured using a 2MP HP 3300 HD Web Camera (1080x1920 pixels) including varying degrees perspective distortion. The DFT was 512x512 and the spacing of cardinal points, $M$, was 256 pixels. Resulting homographies are compared with those obtained using the alternative multi-scale image registration method [2]. The histogram of $E_{\%}$ between the recovered homographies is shown in Figure 9 (right) with median value of 0.32%. A similar comparison was performed using the same parameter settings for 100 simulated distorted images ($R$ = 400) and is shown on the left (median value 0.19%). The small gap between real and simulated data could arise from a number of sources including un-modelled non-linear distortion of the images as a result of lens distortion and motion blur. Accordingly, as the two registration methods operate on slightly different regions of each image their results may diverge slightly.

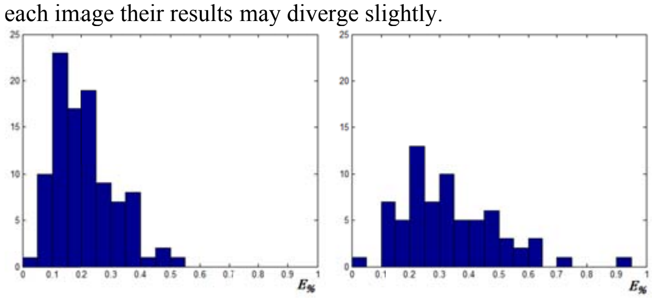


Figure 9. Percentage error histograms for simulated (left) and real (right) data comparing against an image based registration method.

## 4. DISCUSION

We have shown that it is possible to recover planar homographies from clustered-dot halftones by observing the variation in affine approximations based on local Fourier Transforms. As probably expected, the accuracy of the method is not quite as good as that achieved using an existing image registration method. However that method requires that the original image (or halftone) are known to the rectification process and this is not always the case. Advantageously we may wish to embed information in generic halftone images that are not know to the user and this can now be achieved using this novel rectification scheme. The accuracy of the results obtained here suggests that this should be possible but the proof will be in the data recovery rates of the overall approach (this requires that we also recover the remaining translation and will be the emphasis of future work).

Two major choices were made in the design of this method (1) to use the Fourier transform to estimate the local screen directions and (2) to use a synthesized 4 point reconstruction method to estimate the homography based on a pair of estimated north/south cardinal points. The Fourier technique was preferred to the alternative approach of recovering structure directly from the local halftone dot patterns. The local Fourier transform provides an effective, efficient and most importantly robust method for the recovery of the image structure over a region. It also allows us to obtain initial and final (correcting) affine transformations directly from the image. However as the method is an approximation it is necessary to iterate when the distortion is significant.

Likewise other rectifying homographies that leave only affine [11] or similarity [12] can be obtained directly from the data we have extracted (i.e. local screen directions and vanishing points). However we find our approach to be preferable as, provided we start from a reasonable approximation (obtained using AffineDFT), it produces a good solution with roughly the correct scale, orientation and aspect ratio (and only a modest translation). This, amongst other things, makes the task of managing homographies and image rectification much more straight forward.

## 5. REFERENCES


[1] R. Ulichney, M. Gaubatz, S. Simske, "Encoding information in clustered-dot halftones", *Proc. IS&T 26th Intl. Conf. on Digital Printing Technologies and Digital Fabrication (NIP26)*: 602-605. 2010.

[2] S. Pollard, R. Ulichney R, M. Gaubatz and S. Simske, "Forensic Authentication of Data Bearing Halftones", *INSTICC, VISAPP*. 2013.

[3] J-Y. Bouguet, "Pyramid Implementation of Lucas Kanade Feature Tracker: Description of the algorithm", *OpenCV Documents*, Intel Corporation, 1999.

[4] B. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision", *Image Understanding Workshop*, 121-130. 1981.

[5] D.G. Lowe, "Distinctive image features from scale-invariant keypoints", *Intl. Jnl. of Computer Vision,* 60(2): 91-110, 2004.

[6] R.I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd Edition, Cambridge University Press, 2004.

[7] M. Pilu, "Extraction of illusory linear clues in perspectively skewed documents", *IEEE CVPR*: 363–368, 2001

[8] P. Clark, M. Mirmehdi, "Rectifying perspective views of text in 3D scenes using vanishing points", *Pattern Recognition*, 36(11): 2673-2686, 2003.

[9] R.N. Bracewell, K.-Y. Chang, A.K. Jha, Y-H. Wang, "Affine theorem for two-dimensional Fourier transform", *Electronics Letters*, 29(3): 304, 1993,

[10] http://pirsquared.org/research/mcgilldb/welcome.html.

[11] R.T. Collins and J.R. Beveridge. "Matching perspective views of Coplanar structures using projective unwarping and similarity matching". *IEEE CVPR*: 240-245, 1993.

[12] D. Liebowitz and A. Zisserman, "Metric rectification for perspective images of planes", *IEEE CVPR*: 482-488, 1998.