Image Inpainting using Two -View Epipolar Geometry

Prasanna Rangarajan Panos Papamichalis Marc Christensen

Abstract

We take a fresh look at the problem of removing occluders in an image using inpainting. We examine a geometric method that utilizes a second image of the scene from a different viewpoint, to identify the occluded objects. We recover the missing intensities by using the geometric relationship between corresponding points in the two images. The relationship is generally specified by the "epipolar line constraint", and degenerates to a projective homography under special circumstances. For this reason, we use the plane + parallax approach to estimate the epipolar geometry relating the two images. We fill-in missing pixels by copying information from the respective epipolar lines in the second image.

The success of the proposed framework hinges on the ability to estimate the "epipolar line constraint" from noisy correspondences. To this end, we analyze the uncertainty in estimating a homography from noisy correspondences. We rely on the knowledge of this uncertainty to identify parallax vectors best suited for estimating the epipolar geometry.

We do not make any explicit assumptions about the nature or the extent of camera motion, only requiring that the occluded objects are static and undergo limited perspective change. To illustrate the effectiveness of the proposed approach, we "inpaint" occluders in wide-baseline situations.

1. Introduction

The objective of *image inpainting* is to produce a modified image where the inpainted region merges seamlessly with the rest of the image, in a manner not obvious to the untrained eye. Applications of image inpainting include *occluding object removal*, *wire/rig removal*, *retouching of damaged paintings & photographs* and *error concealment*.

Traditional image inpainting methods[1, 2, 5, 12] infer missing intensities using basic principles like continuation of edges (*level lines*), and texture synthesis. The quality of the inpainting depends on the size of the missing region, the geometry of the occluded objects, and the fill-order[1]. Small regions are inpainted effortlessly, while inpainting large regions is difficult and often produces unrealistic results.

The proposed approach extends the scope of current image inpainting methods, by using the *epipolar line constraint* to infer missing intensities from an additional image (*exemplar image*). The use of an *exemplar image* permits us to inpaint large regions, regardless of the geometry of the occluded objects and the fill-order[1].

The use of an *exemplar image* in inpainting is a recent development. Kang et al. [11] recover the missing pixels by assuming that the two images are *affinely* related. Jain[9] assumes that the occluded objects are planar in nature, and recovers the missing intensities using a *projective homography*. Also, there are *spatio-temporal* techniques that attempt to recover missing pixels using several images. Please refer to [15] for a review of such techniques. In this paper, we focus our attention on image inpainting using a single *exemplar image*, for arbitrary camera motion.

The novelty of our approach lies in the use of camera *and-or* occluder motion to reveal the missing intensities. Motion in the occluder either *fully or partially* reveals the occluded objects in the *exemplar image*. In the case of static occluders, camera motion induces an apparent motion in the occluders, depending on their depth. As a result, previously occluded objects are revealed *fully or partially* in the *exemplar image*. In either situation, the occluded objects can experience significant foreshortening due to camera motion. As this makes the task of inpainting difficult, we require that occluded objects undergo limited perspective change.

From now on, we will refer to the *image to be inpainted* as *view-1* and the *exemplar image* as *view-2*.

2. Proposed Approach

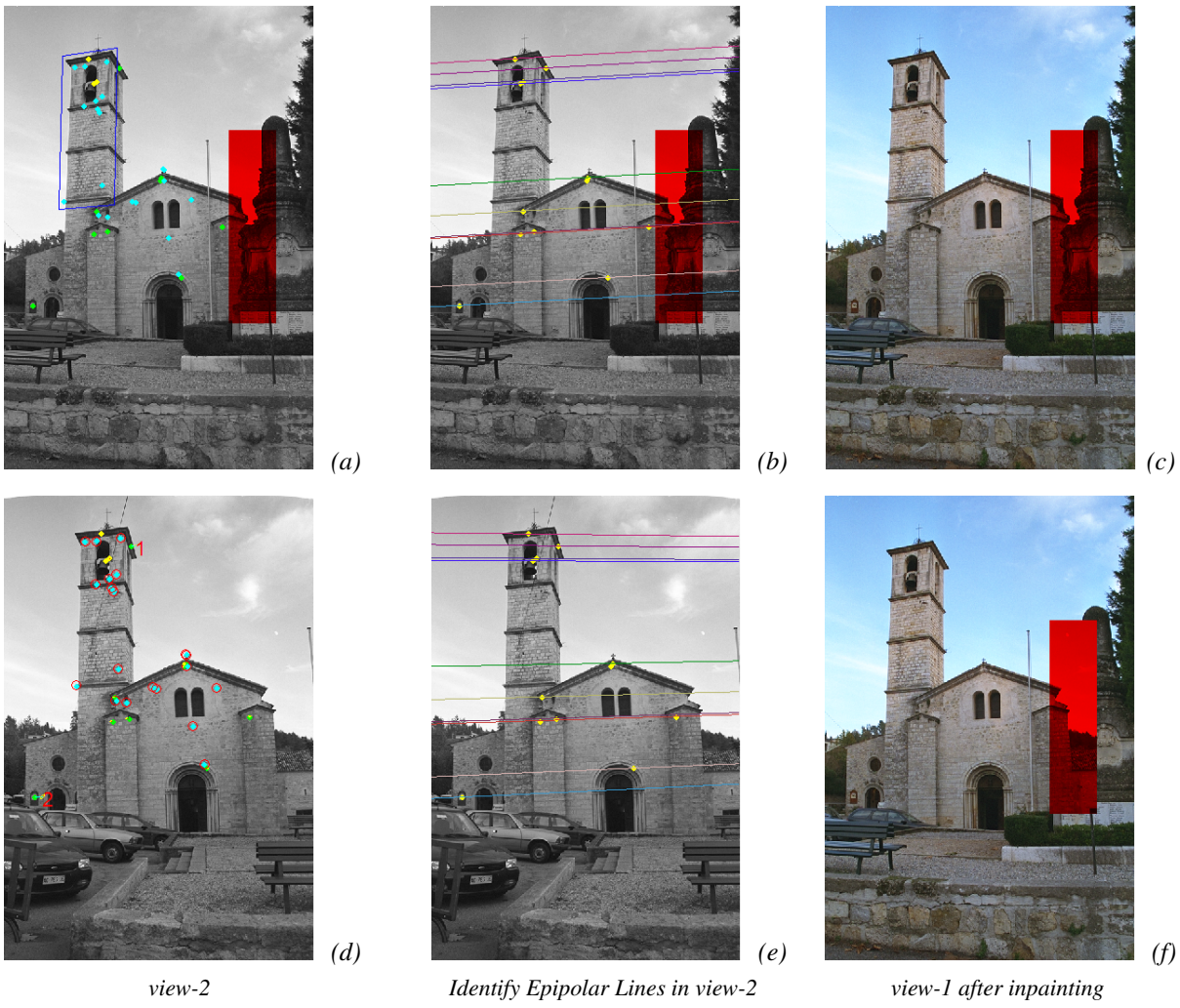
The challenge in the proposed inpainting framework, lies in accurately estimating the geometric relationship between *views-1 & 2* . The relationship is completely specified by a *homography* in these situations

- *occluded object/region belongs to planar facet* **or**
- *small baseline and distant occluders* **or**
- *small change in camera perspective between views-1 & 2*

We inpaint the missing region in *view-1*, by warping *view-2* into *view-1*, using the *inverse homography*. In all other situations, the "*epipolar line constraint*" describes the geometric relationship between corresponding points in *views-1 & 2*. We inpaint missing pixels in *view-1*, by copying intensities from the respective epipolar lines in *view-2*.

In either situation, we need to identify putative point correspondences in the two views. We start by identifying corners with *pixel precision* using the *fast corner detection* algorithm[14], and use *log-polar normalized cross correlation (LP-NCC)*[16] to match the neighborhood of each cor-

[1] fill-order refers to the order in which the missing pixels are inpainted.



The objective is to inpaint parts of the pillar-like structure (transparent red region) in Fig1(a), using an image of the scene (Fig1(d)), from a different viewpoint. We inpaint the missing pixels in Fig1(a) by copying intensities from the respective epipolar lines in Fig1(e). Epipolar lines sharing the same color in Figs1(b),1(e) are in correspondence. The result is shown in Fig1(f). The challenge is to accurately estimate the epipolar geometry relating Fig1(a) and Fig1(d), from a limited number of noisy point correspondences. The proposed approach relies on the uncertainty in estimating homographies and parallax vectors from noisy point correspondences, to accurately estimate the epipolar geometry.

Figure 1. *Image Inpainting using Epipolar Geometry for the Valbonne church scene (courtesy INRIA Robotvis Group)*

ner in *view-1*, with every corner in *view-2*. We prune the list of matches to obtain unique point correspondences. Unfortunately, these point correspondences are noisy. This problem plagues most *corner estimation* and *correspondence matching* algorithms. So, we proceed to estimate the *homography* or the *epipolar geometry* relating the two views, from noisy correspondences (*Steps 1 & 2 in Fig.1*). This is followed by the actual inpainting (*Step 3 in Fig.1*).

The success of the proposed inpainting framework hinges on the ability to estimate the epipolar geometry from noisy correspondences. So we start by analyzing the uncertainty in estimating a homography from noisy correspondences, in Sec–3.2. In Sec–3.3.2, we propagate the uncertainty in estimating a homography to *parallax vectors*. In Sec–4 we examine the use of uncertainty analysis in inpainting, and provide examples of inpainting using the proposed approach in Sec–5.

3. Uncertainty Analysis

Different structures for the uncertainty in a homography, have been proposed by Criminisi[4], Suter[3], Kanatani[10], and Hartley [7]. Unfortunately, these struc-

tures cannot be used directly to find : *parallax vectors* and *planar facets* that are best suited for estimating epipolar geometry from noisy point correspondences. With this objective in mind, we re-examine the problem of determining the uncertainty in a homography. The utility of the proposed structure of uncertainty lies in the ease with which it can be extended to different normalization schemes, and propagated through projective entities such as *algebraic transfer error*, *parallax vectors*, *lines* and *epipoles*.

3.1. Notation

- **boldface** symbols are *vectors*, underlined symbols are *homogeneous vectors*, and CAPITAL letters represent matrices
- The ordered pair (x,y) represents a point in Euclidian 2-space ($\mathbb{R}^2$). The 3-vector $\underline{x} = (x,y,1)^T$ is the corresponding point in *homogeneous coordinates* (Projective 2-space $\mathbb{P}^2$).
- $A(i,:)$ represents the i^{th} row of the matrix A
- $\mathtt{diag}(\mathbf{v})$ is a diagonal matrix with entries of the vector $\mathbf{v}$ along the principal diagonal
- I_d denotes a $d \times d$ identity matrix
- $\mathtt{vec}$-operator[8] vectorizes a matrix by stacking its columns.
- $[\underline{e}]_{\times}\underline{v}$ denotes the use of skew-symmetric matrices to compute the vector product $\underline{e} \times \underline{v}$
- $||\mathbf{v}||$ denotes the $l^2 - norm$ of the vector $\mathbf{v}$

3.2. Homography

A homography is a non-singular matrix that describes a linear mapping in *homogenous coordinates* ($\mathbb{P}^2$). If a homography H_p relates point correspondences $\{(x_i,y_i),(x'_i,y'_i)\}$, the following bilinear relationships hold

$$x'_i = \frac{h_{11}x_i + h_{12}y_i + h_{13}}{h_{31}x_i + h_{32}y_i + h_{33}} \quad y'_i = \frac{h_{21}x_i + h_{22}y_i + h_{23}}{h_{31}x_i + h_{32}y_i + h_{33}} \tag{1}$$

where h_{ij} are entries of the matrix H_p. *eq.(1)* suggests that only the ratio of the entries of H_p is significant. As there are only 8 independent ratios amongst the 9 elements of H_p, a homography has 8 degrees of freedom (d.o.f.). This forms the basis for estimating a homography from 4 or more point correspondences using the Direct Linear Transform algorithm[7].

3.2.1 Direct Linear Transform (DLT) algorithm

Using *kronecker products* and the $\mathtt{vec}$ *operator*[8], we rewrite *eq.(1)* in the form of the linear system

$$\underbrace{\left(\begin{bmatrix} 1 & 0 & -x_i' \\ 0 & 1 & -y_i' \end{bmatrix} \otimes \underline{x_i}^T\right)}_{A_{pi}} \underbrace{\begin{bmatrix} h_{11} \\ h_{12} \\ \cdots \\ h_{33} \end{bmatrix}}_{\mathbf{h_p} \;=\; \mathtt{vec}(H_p^T)} = \underbrace{\begin{bmatrix} 0 \\ 0 \end{bmatrix}}_{\mathbf{0}} \tag{2}$$

Given $M > 4$ sets of point correspondences related by a homography H_p, we estimate H_p by solving the following overdetermined linear system [7]

$$A_p \mathbf{h_p} = \mathbf{0}, \quad A_p = \begin{bmatrix} A_{p1}^T \; \cdots \; A_{pM}^T \end{bmatrix}^T \tag{3}$$

In the absence of noise, A_p is rank-deficient and $\mathbf{h_p} \in$ nullspace of A_p. In the presence of noise A_p is almost always full rank. In either case, $\mathbf{h_p}$ is estimated as the *right singular vector* corresponding to the *least singular value of* A_p[2]. It is easy to observe that $\mathbf{h_p}$ has 8 d.o.f. as $||\mathbf{h_p}|| = 1$.

There is always an uncertainty associated with estimating H_p using DLT, as the point correspondeces used to estimate H_p are noisy. Typically, noise affects corners in both *views-1 & 2*. But, for the sake of analysis we assume that only corners $\{(x'_i,y'_i)\}$ in *view-2* are affected by noise. This assumption is valid only when we seek to find the corner in *view-2* that corresponds to a noisy corner in *view-1*.

3.2.2 Noise model

We assume that the noise $\boldsymbol{\varepsilon}'_{\mathbf{i}} = (\varepsilon'_{xi}, \varepsilon'_{yi})^T$ that affects the corners (x'_i,y'_i) in *view-2* is i.i.d[3] gaussian with zero mean and covariance $E\{\boldsymbol{\varepsilon}'_i \boldsymbol{\varepsilon}'^T_i\} = \mathtt{diag}([\,\sigma_x^2, \sigma_y^2\,])$. This model implies that the error in localizing a corner in the x-direction is independent of the error in the y-direction.

Under these assumptions, we examine the effect of estimating a homography from noisy pixel coordinates, using the DLT algorithm.

3.2.3 Perturbation Analysis of the DLT algorithm

In the presence of noise, we estimate the noisy homography $\mathbf{h_p} + \boldsymbol{\delta}_{h_p}$ as the *right singular vector* corresponding to the *least singular value* of the noisy matrix

$$A_p + N_p = \begin{bmatrix} \begin{bmatrix} 1 & 0 & -(x'_1 + \varepsilon'_{x1}) \\ 0 & 1 & -(y'_1 + \varepsilon'_{y1}) \end{bmatrix} \otimes \underline{x_1}^T \\ \cdots \quad \cdots \quad \cdots \\ \begin{bmatrix} 1 & 0 & -(x'_M + \varepsilon'_{xM}) \\ 0 & 1 & -(y'_M + \varepsilon'_{yM}) \end{bmatrix} \otimes \underline{x_M}^T \end{bmatrix}, \quad \begin{array}{l} \text{where} \\ N_p = [\,\ldots\; {N_p^i}^T \;\ldots\,]^T, \\ N_p^i = \begin{bmatrix} 0 & 0 & -\varepsilon'_{xi} \\ 0 & 0 & -\varepsilon'_{yi} \end{bmatrix} \otimes \underline{x_i}^T \end{array} \tag{4}$$

For small noise levels[4], the effect of estimating a homography from noisy correspondences is best explained by the theory of *first-order perturbation analysis* [4, 6]. Using this theory, we express the perturbation in the homography $\boldsymbol{\delta}_{h_p}$, and its associated uncertainty Λ_{h_p} in the compact form

$$\boldsymbol{\delta}_{h_p} = {A_p}^{\dagger}(W \otimes I_2)\,\Xi \tag{5}$$

$$\Lambda_{h_p} = E\{\boldsymbol{\delta}_{h_p}\boldsymbol{\delta}_{h_p}^T\} = {A_p}^{\dagger}(W \otimes I_2)E\{\Xi\Xi^T\}(W^T \otimes I_2){A_p}^{\dagger T} \tag{6}$$

$$= {A_p}^{\dagger}\left[WW^T \otimes \mathtt{diag}([\,\sigma_x^2, \sigma_y^2\,])\right]{A_p}^{\dagger T} \tag{7}$$

where $A_p^{\dagger}$ is the *pseudoinverse* of the matrix A_p , $W = \mathtt{diag}([\,\omega_1 \;\ldots\; \omega_M\,]) = \mathtt{diag}([\,H_p(3,:)\,\underline{x_1} \;\ldots\; H_p(3,:)\,\underline{x_M}\,])$,

[2] $\mathbf{h_p}$ is the solution to the *Total Least Squares* problem of *eq.(3)*

[3] independent and identically distributed

[4] Quantifying "small" is difficult since it depends on the image size. A noise variance of 4 pixels in a 1024 x 768 image is << noise variance of 4 pixels in a 352 x 288 image.

$\Xi = [\ \varepsilon'_{x1}\ \ \varepsilon'_{y1}\ \ \cdots\ \ \varepsilon'_{xM}\ \ \varepsilon'_{yM}\ ]^T$. The matrix W is comprised of the *homogenous scale factors* $\{\omega_i\}$, while Ξ is comprised of the noise affecting correspondences in *view-2*.

Appendix-1 provides details on deriving the expressions for δ_{h_p} and Λ_{h_p}. Smaller entries in Λ_{h_p} correspond to a smaller uncertainty in estimating $\mathbf{h_p}$, and *vice-versa*.

The proposed structure of Λ_{h_p} (*eq.(7)*) has geometric appeal as explained in Sec-3.2.4, and is easily extended to *normalized DLT*[7] in Sec-3.2.5. In addition, Λ_{h_p} can be propagated to projective entities such as *algebraic transfer error* (Sec-3.2.6), *parallax vectors* (Sec-3.3.2), *lines* and *epipoles*. Further, the structure of Λ_{h_p} in *eq.(7)* allows us to implicitly estimate the unknown noise variances, as described in 3.2.6.

Another advantage of the proposed structure of Λ_{h_p} is that it can be readily expressed in alternative forms[5], such as the structure of the covariance matrix proposed by Hartley [7] and Kanatani [10].

$$\Lambda_{h_p} = \left(\left[(W^{-1} \otimes I_2) A_p \right]^T \left[E\{\Xi\Xi^T\} \right]^{-1} \left[(W^{-1} \otimes I_2) A_p \right] \right)^{\dagger} \tag{8}$$

3.2.4 Structure of the covariance matrix Λ_{h_p}

- Λ_{h_p} is singular, $rank(\Lambda_{h_p}) <= 8$, and $\mathbf{h_p} = \mathtt{null}(\Lambda_{h_p})$
  The redundancy ($||\mathbf{h_p}|| = 1$) *in estimating a homography using DLT, forces* Λ_{h_p} *to be singular. As a result, there is no proper density function associated with the perturbation* δ_{h_p}.
- The uncertainty in $\mathbf{h_p}$ increases with noise variance
  But for large noise variances the expression for Λ_{h_p} *is inaccurate, as first-order perturbation analysis of the matrix* $A_p + N_p$ *is no longer sufficient.*
- The uncertainty in $\mathbf{h_p}$ increases with the *homogenous scale factors* $\{\omega_i\}$
  By choosing point correspondences that are clustered around the fixed-point[6] *of the homography, we can reduce the uncertainty in* $\mathbf{h_p}$. *But, we run the risk of incorrectly estimating a projective transform as being nearly affine. Alternatively, we could arrive at a better estimate of the homography by choosing point correspondences that are spread over a larger area. But this comes at the expense of increased uncertainty (larger* ω_i*'s). This tradeoff is reminiscent of the bias-variance tradeoff in standard estimators.*
- The uncertainty in $\mathbf{h_p}$ depends on the orientation of the observed planar facet
  The least uncertainty in $\mathbf{h_p}$ *results when the camera in view-2 is looking straight at the planar facet (fronto-parallel situation). The reason is that the least error in localizing corners in view-2 results, when camera-2 is in a fronto-parallel situation. Based on this observation, we choose the planar facet that most directly faces camera-2 (reduces uncertainty) and has the largest area (reduces bias), to estimate the epipolar geometry.*

Unfortunately, the DLT algorithm outlined in Sec-3.2.1 cannot be used to estimate the homography H_p, as the ordering of the *right singular vectors of* A_p is sensitive to small perturbations in the pixel coordinates. Hartley[7]points out that the sensitivity is due to the large condition number of the matrix ${A_p}^T A_p$, and observes that an anisotropic normalization of the pixel coordinates alleviates the problem.

Instead of estimating a homography from noisy pixel coordinates, we estimate the homography in normalized coordinates, using the *Normalized DLT algorithm*[7]. To be specific, the normalized homography $\mathbf{h_p^n} + \delta_{h_p^n}$ is estimated as the *right singular vector* corresponding to the *least singular value* of the normalized matrix

$$A_p^n + N_p^n = \begin{bmatrix} \begin{bmatrix} 1 & 0 & -(x_1^{n\prime} + \varepsilon_{x1}^{n\prime}) \\ 0 & 1 & -(y_1^{n\prime} + \varepsilon_{y1}^{n\prime}) \end{bmatrix} \otimes \underline{x_1^n}^T \\ \cdots \\ \begin{bmatrix} 1 & 0 & -(x_M^{n\prime} + \varepsilon_{xM}^{n\prime}) \\ 0 & 1 & -(y_M^{n\prime} + \varepsilon_{yM}^{n\prime}) \end{bmatrix} \otimes \underline{x_M^n}^T \end{bmatrix} \tag{9}$$

$$\underline{x_i^n} = \frac{N_1}{N_1(3,3)} \underline{x_i}\,, \begin{bmatrix} x_i^{n\prime} + \varepsilon_{xi}^{n\prime} \\ y_i^{n\prime} + \varepsilon_{yi}^{n\prime} \\ 1 \end{bmatrix} = \frac{N_2}{N_2(3,3)} \underbrace{\begin{bmatrix} x_i' + \varepsilon_{xi}' \\ y_i' + \varepsilon_{yi}' \\ 1 \end{bmatrix}}_{\widehat{\underline{x_i'}}} \tag{10}$$

$$N_1 = \mathtt{ichol}\left(\sum_{i=1}^{M} \underline{x_i}\,\underline{x_i}^T \right), N_2 = \mathtt{ichol}\left(\sum_{i=1}^{M} \widehat{\underline{x_i'}}\,\widehat{\underline{x_i'}}^T \right) \tag{11}$$

The *inverse cholesky factors*[6] N_1 & N_2 normalize pixel coordinates in *views-1 & 2* respectively. For small noise levels, the dependence of N_2 on the noise terms in *eq.(11)* is so negligible that we ignore it. The homography H_p in pixel coordinates is related to the homography H_p^n in normalized coordinates as

$$H_p = \frac{N_2^{-1}\ H_p^n\ N_1}{||N_2^{-1}\ H_p^n\ N_1||},\ \mathbf{h_p} = \frac{(N_2^{-1} \otimes N_1^T)\ \mathbf{h_p^n}}{||(N_2^{-1} \otimes N_1^T)\ \mathbf{h_p^n}||} = \frac{\mathbf{h}}{||\mathbf{h}||} \tag{12}$$

3.2.5 Perturbation Analysis of Normalized DLT

Along the lines of Sec-3.2.3, we can express the perturbation in the normalized homography ($\delta_{h_p^n}$) as

$$\delta_{h_p^n} = {A_p^n}^{\dagger} (W^n \otimes I_2)\, \Xi^n \tag{13}$$

$W^n = \mathtt{diag}([\ H_p^n(3,:)\,\underline{x_1^n} \ldots H_p^n(3,:)\,\underline{x_M^n}\ ])$, $\Xi^n = [\ \varepsilon_{x1}^{n\prime}\ \ \varepsilon_{y1}^{n\prime} \ldots \varepsilon_{xM}^{n\prime}\ \ \varepsilon_{yM}^{n\prime}\ ]^T$
The corresponding uncertainty is specified as

$$\Lambda_{h_p^n} = {A_p^n}^{\dagger} \left[W^n {W^n}^T \otimes N_{2S}\ \mathtt{diag}([\ \sigma_x^2, \sigma_y^2\ ])\, N_{2S}^T \right] {A_p^n}^{\dagger T} \tag{14}$$

$$\text{where}\quad N_{2S} = \begin{bmatrix} N_2(1,1) & N_2(1,2) \\ N_2(2,1) & N_2(2,2) \end{bmatrix}$$

Appendix-2 provides details on deriving $\Lambda_{h_p^n}$. Notice that the structure of $\Lambda_{h_p^n}$ in *eq.(14)* is similar to the structure of Λ_{h_p} in *eq.(7)*, with the exception of the term N_{2S}.

[5] Using the property $(A\,B)^{\dagger} = B^{\dagger}\,A^{\dagger}$, we can rewrite *eq.(6)* as *eq.(8)*

[6] $\underline{x_i}' = \underline{x_i}, \quad \Rightarrow \omega_i = 1$

However, *eq.(14)* only describes the uncertainty of the normalized homography $\mathbf{h_p^n}$. We need to propagate $\Lambda_{h_p^n}$ through *eq.(12)*, to determine Λ_{h_p}, in pixel coordinates. The process begins by propagating noise through $\mathbf{h}$ in *eq.(12)*, and finally through the normalization $\frac{\mathbf{h}}{||\mathbf{h}||}$. The uncertainty in $\mathbf{h_p}$ has the following structure

$$\Lambda_{h_p} = \frac{1}{||\mathbf{h}||}\left[I_9 - \frac{\mathbf{h}\mathbf{h}^T}{||\mathbf{h}||^2}\right]\Lambda_h\left[I_9 - \frac{\mathbf{h}\mathbf{h}^T}{||\mathbf{h}||^2}\right]\frac{1}{||\mathbf{h}||} \tag{15}$$

where $\mathbf{h} = (N_2^{-1}\otimes N_1^T)\mathbf{h_p^n}$,and $\Lambda_h = (N_2^{-1}\otimes N_1^T)\ \Lambda_{h_p^n}$[7] $(N_2^{-T}\otimes N_1)$

Substituting for $\Lambda_{h_p^n}$ from *eq.(14)*, we find that

$$\boxed{\begin{aligned}\Lambda_{h_p} &= \mathscr{M}_h A_p^{n\dagger}\left[(W^n)^2\otimes N_{2S}\,\mathtt{diag}([\,\sigma_x^2,\,\sigma_y^2\,])\,N_{2S}^T\right]A_p^{n\dagger T}\mathscr{M}_h{}^T\\ \text{where } \mathscr{M}_h &= \frac{1}{||\mathbf{h}||}\left[I_9 - \frac{\mathbf{h}\mathbf{h}^T}{||\mathbf{h}||^2}\right]\left(N_2^{-1}\otimes N_1^T\right)\end{aligned}} \tag{16}$$

Alternatively, we can rewrite *eq.(16)* in terms of σ_x^2 & σ_y^2 as

$$\begin{aligned}\Lambda_{h_p} &= \sigma_x^2\ \Lambda_{h_p}^x\ +\ \sigma_y^2\ \Lambda_{h_p}^y \qquad (17)\\ \Lambda_{h_p}^x &= \mathscr{M}_h A_p^{n\dagger}\left[(W^n)^2\otimes N_{2S}\,\mathtt{diag}([\,1,\,0\,])N_{2S}^T\right]A_p^{n\dagger}\mathscr{M}_h{}^T\\ \Lambda_{h_p}^y &= \mathscr{M}_h A_p^{n\dagger}\left[(W^n)^2\otimes N_{2S}\,\mathtt{diag}([\,0,\,1\,])N_{2S}^T\right]A_p^{n\dagger}\mathscr{M}_h{}^T\end{aligned}$$

In Sec-3.2.6, we use this structure of Λ_{h_p} to find σ_x^2 & σ_y^2. Comments about the structure of the uncertainty Λ_{h_p} from Sec-3.2.4 are still valid.

We conclude the study of uncertainty in estimating homographies using DLT, by noting that the estimated homography $\widehat{\mathbf{h_p}}$ is a random vector with

mean	$\mathbf{h_p}$	*true homography defined in eq.(12)*
covariance	Λ_{h_p}	*defined in eq.(16)*

Note: There is *no* proper density function associated with $\widehat{\mathbf{h_p}}$, as Λ_{h_p} is singular.

3.2.6 Numerically estimating Λ_{h_p}

Provided we know the noise variances (σ_x^2, σ_y^2), we can find an estimate $\widehat{\Lambda_{h_p}}$ of the true covariance Λ_{h_p}, by replacing *true* quantities (A_p^n, H_p^n) in *eq.(17)*, with *observed or estimated* quantities ($A_p^n + N_p^n, \widehat{H_p^n}$).

In the absence of noise statistics, we estimate σ_x^2, σ_y^2 from the statistics of the *algebraic transfer error* in *view-2*. The *algebraic transfer error* for the i^{th} point correspondence is defined as

$$\mathbf{r_i'} = \begin{bmatrix} x_i' + \varepsilon_{xi}' \\ y_i' + \varepsilon_{yi}' \end{bmatrix} - \underbrace{\frac{1}{\widehat{H_p}(3,:)\,\underline{x_i}}\begin{bmatrix}\widehat{H_p}(1,:)\,\underline{x_i}\\ \widehat{H_p}(2,:)\,\underline{x_i}\end{bmatrix}}_{\mathbf{y_i'}} \tag{18}$$

[7] From *eq.(12)* we find that a perturbation $\delta_{h_p}^n$ in $\mathbf{h_p^n}$, induces the perturbation $\delta_h = (N_2^{-1}\otimes N_1^T)\,\delta_{h_p^n}$, in $\mathbf{h}$. The corresponding uncertainty is $\Lambda_h = E\{\delta_h\delta_h{}^T\} = (N_2^{-1}\otimes N_1^T)\ \Lambda_{h_p^n}\ (N_2^{-T}\otimes N_1)$

$\mathbf{r_i'}$ is a random vector with 2 sources of uncertainty

1. noisy point correspondences in view-2

2. error in point transfer using the noisy homography $\widehat{H_p}$

Assuming the *transfer errors* to be gaussian distributed[8],

$$\mathbf{r_i'} \sim \mathscr{N}\left(\mathbf{0}, \mathtt{diag}([\,\sigma_x^2, \sigma_y^2\,]) + \Lambda_{\mathbf{y_i'}}\right),\quad \Lambda_{\mathbf{y_i'}} = J_{h_p}^i \Lambda_{h_p} J_{h_p}^{i\ T} \tag{19}$$

$$J_{h_p}^i = \frac{1}{H_p(3,:)\,\underline{x_i}}\begin{bmatrix} H_p(3,:)\,\underline{x_i} & 0 & -H_p(1,:)\,\underline{x_i} \\ H_p(3,:)\,\underline{x_i} & 0 & -H_p(2,:)\,\underline{x_i}\end{bmatrix} \otimes \underline{x_i}^T \tag{20}$$

Please refer to [7] for details on deriving $\Lambda_{\mathbf{y_i'}}$. It can be shown that the matrix $\mathrm{R} = [\,\mathbf{r_1'}\ \ldots\ \mathbf{r_M'}\,]$ formed by stacking the *algebraic transfer errors* for the M point correspondences, has the following statistics

$$\begin{aligned}E\{\mathrm{R}\mathrm{R}^T\} &= \sum_{i=1}^{M} E\{\mathbf{r_i'}\mathbf{r_i'}^T\} = M\begin{bmatrix}\sigma_x^2 & 0\\ 0 & \sigma_y^2\end{bmatrix} + \sum_{i=1}^{M} J_{h_p}^i \Lambda_{h_p} J_{h_p}^{i\ T}\\ &= M\begin{bmatrix}\sigma_x^2 & 0\\ 0 & \sigma_y^2\end{bmatrix} + \sum_{i=1}^{M}\left[\sigma_x^2 J_{h_p}^i \Lambda_{h_p}^x J_{h_p}^{i\ T} + \sigma_y^2 J_{h_p}^i \Lambda_{h_p}^y J_{h_p}^{i\ T}\right] \qquad (21)\end{aligned}$$

The *left-hand-side* of *eq.(21)* can be estimated directly from the data. We find estimates $\widehat{\sigma_x^2}, \widehat{\sigma_y^2}$ of the true noise variances σ_x^2, σ_y^2, by solving the linear system of *eq.(21)*, after these substitutions : $\widehat{J_{h_p}^i}, \widehat{\Lambda_{h_p}^x}, \widehat{\Lambda_{h_p}^y} \rightarrow J_{h_p}^i, \Lambda_{h_p}^x, \Lambda_{h_p}^y$.

3.2.7 Refining the homography $\widehat{H_p}$

We now examine how knowledge of $\widehat{\Lambda_{h_p}}$ could be used to refine $\widehat{H_p}$. Suppose, $\mathscr{S}$ is the set of noisy correpondences in the two views, and $\mathscr{S}_{h_p}$ is the subset with homography $\widehat{H_P}$. We refine $\widehat{H_p}$ by using correspondences in $\mathscr{S}\backslash\mathscr{S}_{h_p}$[9] that are *consistent* with $\widehat{H_p}$, while discarding *inconsistent* correspondences in $\mathscr{S}_{h_p}$. We declare the i^{th} point correspondence as *consistent* only when the *Mahalanobis norm* of its *algebraic transfer error* given by $d_{alg}[i]$ is < 5.9915.

$$d_{alg}[i] = \mathbf{r_i'}^T\left(\begin{bmatrix}\widehat{\sigma_x}^2 & 0\\ 0 & \widehat{\sigma_y}^2\end{bmatrix} + \widehat{J_{h_p}}\widehat{\Lambda_{h_p}}\widehat{J_{h_p}}^T\right)\mathbf{r_i'} \tag{22}$$

where $\widehat{\sigma_x}^2, \widehat{\sigma_y}^2$ are the estimated noise variances, and $\widehat{J_{h_p}}$ is defined in *eq.(20)*. The number 5.9915 specifies the $\mathtt{Probability}(\textit{misclassifying an inconsistent correspondence}) < 5\%$[7]. This follows from the fact that $d_{alg}[i]$ is a χ_2^2 distributed random variable with 2 d.o.f. The proposed thresholding scheme is based on the thresholding scheme for *sampson distances*, utilized in estimating homographies using RANSAC[7].

The success of the proposed inpainting scheme depends on the ability to reliably estimate the epipolar geometry from noisy correspondences. So we examine the effect of noise on estimating the *epipolar line constraint*.

[8] The gaussian assumption breaks down in the presence of outliers.

[9] The set difference $A\backslash B$ is defined as $A\backslash B = \{x : x \in A\ \ \&\ \ x \notin B\}$

3.3. Epipolar Geometry

Given two views of a static scene, the *epipolar line constraint*[7] describes a generic relationship between corresponding points[10] in the two views. Suppose $\underline{x_j} = [x_j, y_j, 1]^T$ and $\underline{x'_j} = [x'_j, y'_j, 1]^T$ are corresponding points[10] in *views-1 & 2* respectively. They satisfy the *epipolar line constraint*

$$\underline{x'_j}^T \underbrace{F\underline{x_j}}_{l_{j2}} = 0 \quad , \quad \underline{x_j}^T \underbrace{F^T\underline{x'_j}}_{l_{j1}} = 0 \tag{23}$$

where F is a singular matrix depending on the relative position and orientation of the cameras in the two views. The matrix F is referred to as the fundamental matrix. The *epipolar line constraint* derives its name from the fact that it describes a point-line incidence relation[11]. The entities l_{j1}, l_{j2} are referred to as conjugate epipolar lines.

> If $\underline{x_j}$ is a missing pixel, its corresponding point in *view-2* must satisfy the *epipolar line constraint* of *eq.(23)*. We exploit this knowledge to inpaint the missing pixel $\underline{x_j}$.

3.3.1 Estimating the Fundamental matrix

Given a planar facet that induces a homography H_p between the two views and the epipole[12] $\underline{e'}$ in *view-2*, the fundamental matrix F is specified as

$$F = [\underline{e'}]_{\mathbf{x}} H_p \tag{24}$$

The epipole $\underline{e'}$ is the intersection of all *epipolar lines* in *view-2*. As mentioned before, if $\underline{x_j}, \underline{x'_j}$ are corresponding points in the two views, $\underline{x'_j}$ lies on the line $l_{j2} = F\underline{x_j}$. It can be shown[7] that the point $\underline{\widetilde{x}'_j} \sim H_p \underline{x_j}$ [13], also lies on the line l_{j2}. This means that the line joining $\underline{x'_j}$ and $\underline{\widetilde{x}'_j}$ is a fragment of the epipolar line l_{j2}. Given two such fragments we can compute $\underline{e'}$, and the full epipolar geometry follows from *eq.(24)*. The vector joining $\underline{\widetilde{x}'_j}$ and $\underline{x'_j}$ is referred to as the *parallax vector*. Consequently, this approach to estimating F is referred to as the *plane+parallax* method.

To reliably estimate the fundamental matrix using the *plane+parallax* method, we need to identify atleast two parallax vectors with little-or-no uncertainty in their orientation. This is a challenging task given noisy point correspondences in the two views. For short parallax vectors ($\underline{x'_j}$ & $\underline{\widetilde{x}'_j}$ are close), even small perturbations in the end-points significantly alters the orientation of the parallax vector. This suggests that the length of a parallax vector could be used to determine its reliability. But, it is likely that grossly mismatched correspondences are mistaken for reliable parallax vectors. So, we assume that grossly mismatched correspondences have been removed prior to inpainting. The sources of error affecting the orientation of a parallax vector are

1. *noise in the correspondence* $\underline{x_j}'$
2. *noise in* $\underline{\widetilde{x}_j}'$ *due to noise in the estimated homography* $\widehat{H_p}$

We now examine how uncertainty in estimating H_p can be combined with the length of a parallax vector, to identify the two most reliable *parallax vectors*.

3.3.2 Reliability of *parallax vectors*

Given a noisy homography $\widehat{H_p}$, an estimate of its uncertainty $\widehat{\Lambda_{h_p}}$, noise variances $\widehat{\sigma_x^2}, \widehat{\sigma_y^2}$, and a noisy point correspondence $\{(x_j, y_j), (x'_j + \varepsilon'_{xj}, y'_j + \varepsilon'_{yj})\}$ that is *inconsistent*[14] with the homography, we can construct an epipolar line segment by joining $\left(\frac{\widehat{H_p}(1,:)\underline{x_j}}{\widehat{H_p}(3,:)\underline{x_j}}, \frac{\widehat{H_p}(2,:)\underline{x_j}}{\widehat{H_p}(3,:)\underline{x_j}} \right)$ and $(x'_j + \varepsilon'_{xj}, y'_j + \varepsilon'_{yj})$. The length of this line segment is

$$d_j = \left\| \begin{bmatrix} x'_j + \varepsilon'_{xj} \\ y'_j + \varepsilon'_{yj} \end{bmatrix} - \frac{1}{\widehat{H_p}(3,:)\underline{x_j}} \begin{bmatrix} \widehat{H_p}(1,:)\underline{x_j} \\ \widehat{H_p}(2,:)\underline{x_j} \end{bmatrix} \right\|^2 = ||\mathbf{r_j}||^2 \tag{25}$$

It is easy to observe that d_j is a random variable. Using the noise model of Sec-3.2.2 and the covariance structure ($\Lambda_{\mathbf{y_j}'}$) of $\mathbf{y'_j}$ defined in *eqs.(19) and (20)*, we can prove that

$$E\{d_j\} = \widetilde{d}_j + trace\left(\begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix} \right) + trace\left(\Lambda_{\mathbf{y'_j}} \right)$$

$$\widetilde{d}_j = \left\| \begin{bmatrix} x'_j \\ y'_j \end{bmatrix} - \frac{1}{H_p(3,:)\underline{x_j}} \begin{bmatrix} H_p(1,:)\underline{x_j} \\ H_p(2,:)\underline{x_j} \end{bmatrix} \right\|^2 \tag{26}$$

where $\widetilde{d}_j$ is the true length of the epipolar line segment. We define the reliability of a parallax vector as the ratio $\widetilde{d}_j^{-1}(\sigma_x^2 + \sigma_y^2 + trace(\Lambda_{\mathbf{y'_j}}))$. The ratio is minimum when the noise is minimum and the length of the line segment is maximum. Given 2 or more correspondences that are inconsistent with the homography $\widehat{H_p}$, we find the two *parallax vectors* with the highest reliability in the sense of

$$d_j^{-1}\left(\widehat{\sigma_x}^2 + \widehat{\sigma_y}^2 + trace(\widehat{J_{h_p}}\widehat{\Lambda_{h_p}}\widehat{J_{h_p}}^T) \right) \tag{27}$$

where d_j is the observed length of the epipolar line segment, defined in *eq.(25)*. The chosen *parallax vectors* are intersected to find the epipole $\underline{e'}$ in *view-2*. The full epipolar geometry follows from *eq.(24)*.

4. Putting it all together

Fig.(1) summarizes the process of inpainting in the proposed approach. The user first selects images corresponding to *views-1 & 2*. We then procced to identify putative correspondences in the two views (*colored dots in Figs1(a),1(d)*).

[10] images of the same 3D world point in the two views

[11] (x_j, y_j) lies on line l_{j1} in *view-1*, and (x'_j, y'_j) lies on line l_{m2} in *view-2*

[12] The epipole $\underline{e'}$ is the image of the 1^{st}*-pinhole* as seen by *camera-2*.

[13] $\underline{\widetilde{x}_j}'$ represents the point $\underline{x_j}$ transferred by the homography H_p

[14] $d_{alg}[j] > 5.9915$, where $d_{alg}[j]$ is defined in *eq.(22)*

Next, the user outlines *a hypothetical plane $\mathscr{R}$ & the regions to be inpainted* $\{\Omega_k\}$, in *view-1*. The *hypothetical plane* outlined in *blue* in *Fig1(a)*, is any region that includes correspondences from a planar facet. Using the set of correspondences $\in \mathscr{R}$, we estimate a homography $\widehat{H_p}$, its associated uncertainty $\widehat{\Lambda_{h_p}}$, and noise variances $\widehat{\sigma_x^2}, \widehat{\sigma_y^2}$. Using the strategy of Sec-3.2.7, we refine the homography $\widehat{H_p}$ until the ratio $\frac{\widehat{\sigma_x}^2+\widehat{\sigma_y}^2}{\#consistent\ \ correspondences}$ is a minimum. The result of the refinement is shown in *Figs.1(a),1(d)*. The cyan points in *Figs.1(a),1(d)* are *consistent correspondences*, while the yellow & green points are the *inconsistent correspondences* $\in \mathscr{R}$ & $\notin \mathscr{R}$ respectively. We use the refined homography $\widehat{H_p}$ to determine the two most reliable parallax vectors, as outlined in Sec-3.3.2. The best set of parallax vectors labeled 1 and 2 in *Fig.1(d)* are intersected to find the epipole $\underline{e'}$ in *view-2*. The full epipolar geometry follows from *eq.(24)*.
If there is more than one planar facet in the scene, the best inpainting is achieved for the planar facet that most directly faces *camera-2*, as described in Sec-3.2.4.

4.1. Inpainting using the "*epipolar line constraint*"

From Sec-3.3, we know that if $\underline{x_m}$ is a missing pixel in *view-1*, its corresponding point in *view-2* satisfies the *epipolar line constraint* of *eq.(23)*. Suppose $l_{m1} = [\mathtt{null}(F)]_{\mathtt{x}}\ \underline{x_m}$, and $l_{m2} = F\underline{x_m}$ are conjugate epipolar lines in *views-1 & 2*, corresponding to $\underline{x_m}$. In principle, we could inpaint $\underline{x_m}$ by searching for the best matching patch along the epipolar line l_{m2}, in *view-2*. But this is time consuming, computationally intensive, and not always correct. A simpler alternative to inferring missing pixels is globally strechning/shrinking l_{m2} until fragments of l_{m1}&l_{m2} are aligned.

We repeat the steps outlined below for each pair of conjugate epipolar lines $\{l_{m1}, l_{m2}\}$, in each region Ω_k.

1. Identify landmarks on l_{m1} & l_{m2} as points with the largest gradient magnitude along the epipolar line ($\Rightarrow$ *edge runs across the epipolar line*)
2. Match the neighborhood around each landmark in *view-1* with every landmark in *view-2*, using *LP-NCC*
3. Find atleast 2 uniquely matching landmarks on l_{m1} & l_{m2}
4. Rotate l_{m1} & l_{m2} around the respective epipoles
5. Find the 1-d affine transform A_m that alings l_{m1} & l_{m2}, using the coordinates of the matching landmarks
6. Transform the points $(x_m, y_m) \in (l_{m1} \cap \Omega_k)$ using A_m, to find points on l_{m2} with the missing intensities

Any error in identifying landmarks introduces errors in the 1-d affine transform and hence the inpainting. Failure to identify landmarks results in the inability to inpaint a specific conjugate line pair.

Contrary to our assumption, conjugate epipolar lines can be aligned using a *locally affine* transform. We are currently investigating the use of *elastic registration*[13] techniques to locally align a pair of conjugate epipolar lines.

5. Results

The effectiveness of the proposed approach is best demonstrated in *Fig.1*, where we attempt to inpaint parts of a pillar like structure. It is difficult to accurately estimate the epipolar geometry in this example, as we have a limited number of correspondences, and short parallax vectors. A quick look at *Fig.1(d)* reveals that the proposed inpainting scheme has succeeded in inpainting the pillar like structure. The quality of the inpainting can be judged by the accuracy with which corners on the roof, and vertical structures within the missing region have been reconstructed.

We provide two more examples of the proposed inpainting, wherein traditional image inpainting methods would have failed. In the first example (*Figs.2(a)-2(d)*) we attempt to recover the missing detail in the bookshelf, while in the second example (*Figs.2(e)-2(h)*) we inpaint a lamp & the window panes behind it. A quick look at *Fig.2(a),2(h)*, reveals that we have succesully recovered the missing detail in the bookshelf of *Fig.2(a)*, and inpainted the lamp & the missing window panes in *Fig.2(h)*. Notice that we succeed in recovering the arch like structure on the window panes, in *Fig.2(h)*. But, a closer examination of *Figs.2(d),2(h)* reveals that the some of the objects are scaled incorrectly. This is a drawback of assuming that a pair of conjugate epipolar lines can be aligned globally using a 1-d affine transform.

6. Conclusions

We extend the scope of image inpainting by using the *epipolar line constraint* to infer the missing pixels from an *exemplar image*. We overcome the difficulty in accurately estimating the epipolar geometry from noisy point correspondences, by using the *proposed structure of uncertainty in a homography* and the *proposed measure of the reliability of a parallax vector*.

The quality of the inpainting can be improved further, by
1. *using elastic matching to align conjugate epipolar lines*
2. *non-linear refinement*[7] of the estimated homography
3. *using additional views to fully inpaint the occluding objects.*

Appendix-1. Perturbation in DLT homography

Using *first-order perturbation analysis*[6] we can express the perturbation in the DLT homography (δ_{h_p}) in terms of the *unperturbed singular values & right singular vectors* $\{\sigma_j, \mathbf{v_j}\}_{j=1}^{9}$ of A_p.

$$\delta_{h_p} = -\sum_{j=1}^{8}\left(\frac{\mathbf{v_j}\mathbf{v_j}^T}{\sigma_j^2}\left[A_p^T N_p \mathbf{h_p} + N_p^T A_p \mathbf{h_p}\right]\right) \qquad \text{(A1)}$$

We simplify *eq.(A1)* by noting that $A_p\mathbf{h_p} = \mathbf{0}$, and
- *o* $A_p^T = \sum_{k=1}^{9}\sigma_k \mathbf{v_k}\mathbf{u_k}^T$ *singular value decomposition of* A_p
- *o* $\mathbf{v_m}^T\mathbf{v_n} = \delta[m-n]$ *orthogonality of right singular vectors*

$$\delta_{h_p} = -\left(\sum_{j=1}^{8}\frac{\mathbf{v_j}\mathbf{u_j}^T}{\sigma_j}\right)N_p\mathbf{h_p} = -A_p^{\dagger}\ N_p\mathbf{h_p} \qquad \text{(A2)}$$

where $A_p^{\dagger}$ is the *pseudoinverse* [6] of the matrix A_p.
Using *eq.(4)* we can express the term $N_p\mathbf{h_p}$ as

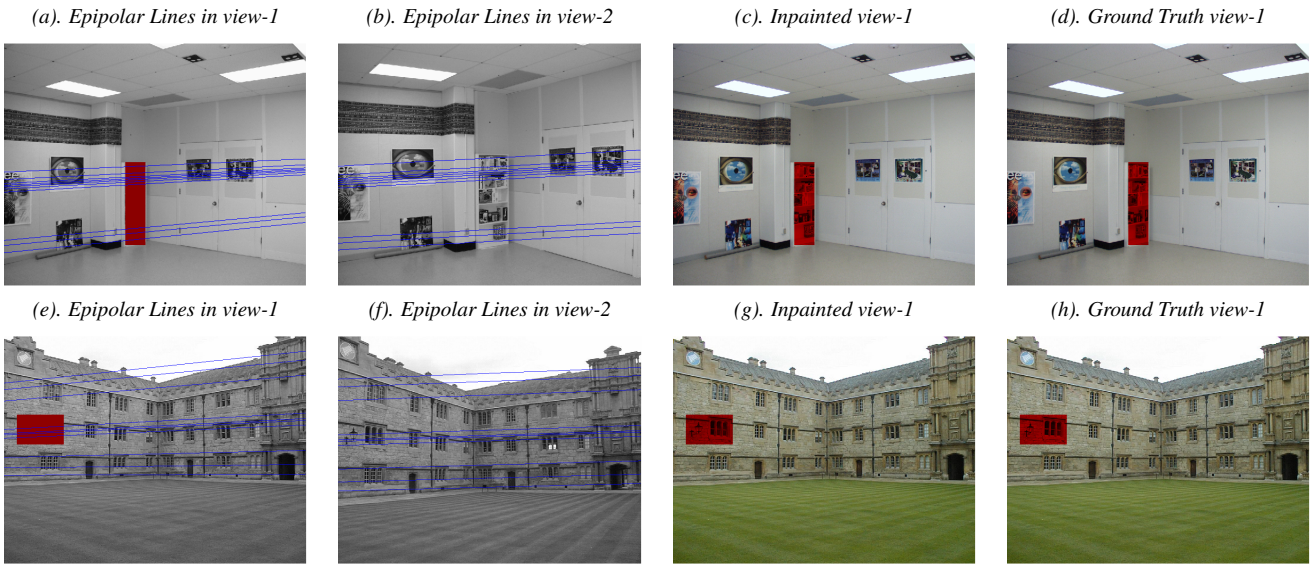


Figure 2. *Image Inpainting for an indoor scene (courtesy NRC-IIT VIT Group), and an outdoor scene (courtesy Oxford Univ. VGG)*

$$N_p \mathbf{h_p} = \begin{bmatrix} \left\{ \begin{bmatrix} 0 & 0 & -\varepsilon'_{x1} \\ 0 & 0 & -\varepsilon'_{y1} \end{bmatrix} \otimes \underline{x_1}^T \right\} \mathtt{vec}(H_p{}^T) \\ \cdots \\ \left\{ \begin{bmatrix} 0 & 0 & -\varepsilon'_{xm} \\ 0 & 0 & -\varepsilon'_{ym} \end{bmatrix} \otimes \underline{x_M}^T \right\} \mathtt{vec}(H_p{}^T) \end{bmatrix} \tag{A3}$$

Using the property $\mathtt{vec}(A\ B\ C) = (C^T \otimes A)\ \mathtt{vec}(B)$, we find that

$$N_p \mathbf{h_p} = \begin{bmatrix} \mathtt{vec}\left(\underline{x_1}^T H_p^T \begin{bmatrix} 0 & 0 & -\varepsilon'_{x1} \\ 0 & 0 & -\varepsilon'_{y1} \end{bmatrix}^T \right) \\ \cdots\cdots \\ \mathtt{vec}\left(\underline{x_M}^T H_p^T \begin{bmatrix} 0 & 0 & -\varepsilon'_{xM} \\ 0 & 0 & -\varepsilon'_{yM} \end{bmatrix}^T \right) \end{bmatrix} = (W \otimes I_2)\ \Xi \tag{A4}$$

$$W = \mathtt{diag}([\ H_p(3,:)\,\underline{x_1} \ldots H_p(3,:)\,\underline{x_M}\ ])\,,\ \Xi = [\ \varepsilon'_{x1}\ \ \varepsilon'_{y1}\ \ \ldots\ \ \varepsilon'_{xM}\ \ \varepsilon'_{yM}\ ]^T$$

Using *eq.(A4)* in *eq.(A2)*, we can expres δ_{h_p} in compact form as

$$\delta_{h_p} = A_p{}^{\dagger}\,(W \otimes I_2)\,\Xi \tag{A5}$$

The uncertainty in the DLT homography is estimated as

$$\Lambda_{h_p} = E\{\delta_{h_p}\delta_{h_p}{}^T\} = A_p{}^{\dagger}\,(W \otimes I_2)\,E\{\Xi\Xi^T\}(W^T \otimes I_2)A_p{}^{\dagger T} \tag{A6}$$

Using the noise model of 3.2.2, we can show that

$$E\{\Xi\Xi^T\} = I_M \otimes \mathtt{diag}([\ \sigma_x^2,\ \sigma_y^2\ ]) \tag{A7}$$

Substituting *eq.(A7)* in *eq.(A6)*, and using the property $(A \otimes B)(C \otimes D) = (A\ B \otimes C\ D)$, we obtain the expression for Λ_{h_p} (*eq.(7)*) in compact form.

Appendix-2. Uncertainty in Normalized DLT

From *eq.(13)*, the perturbation in the normalized homography is $\delta_{h_p^n} = A_p^{n\,\dagger}\,(W^n \otimes I_2)\,\Xi^n$. The corresponding uncertainty $\Lambda_{h_p^n}$ is

$$\Lambda_{h_p^n} = E\left\{\delta_{h_p^n}\delta_{h_p^n}{}^T\right\} = A_p^{n\,\dagger}\,(W^n \otimes I_2)\,E\{\Xi^n\Xi^{nT}\}(W^{nT} \otimes I_2)\ A_p^{n\,\dagger T} \tag{A8}$$

To model the statistics of the noise vector Ξ^n , we need to examine the effect of normalizing the noise terms $\{(\varepsilon'_{xi}, \varepsilon'_{yi})\}$. From *eq.(10)*,

$$\begin{bmatrix} \varepsilon_{xi}^{n'} \\ \varepsilon_{yi}^{n'} \end{bmatrix} = \begin{bmatrix} N_2(1,1) & N_2(1,2) \\ N_2(2,1) & N_2(2,2) \end{bmatrix} \begin{bmatrix} \varepsilon'_{xi} \\ \varepsilon'_{yi} \end{bmatrix} = N_{2S} \begin{bmatrix} \varepsilon'_{xi} \\ \varepsilon'_{yi} \end{bmatrix} \tag{A9}$$

$$\Rightarrow \Xi^n = [\ \varepsilon_{x1}^{n'}\ \ \varepsilon_{y1}^{n'}\ \ \ldots\ \ \varepsilon_{xM}^{n'}\ \ \varepsilon_{yM}^{n'}\ ]^T = (I_M \otimes N_{2S})\,\Xi \tag{A10}$$

Using *eqs.(A7,A10)* in *eq.(A8)* and simplifying, we obtain the expression (*eq.(14)*) for $\Lambda_{h_p}^n$ in compact form.

References

[1] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester. Image inpainting. Siggraph 2000, Computer Graphics Proceedings, pages 417–424, 2000. 1

[2] T. Chan and J. Shen. Non-texture inpaintings by curvature-driven diffusions. J. Visual Communication and Image Representation, 12:436–449, 2001. 1

[3] P. Chen and D. Suter. Homography estimation and heteroscedastic noise - a first order perturbation analysis, Technical Report MECSE-32-2005. 2

[4] A. Criminisi. Modelling and using uncertainties in video metrology, Technical Report, 1997. 2, 3

[5] A. Criminisi. Object removal by exemplar-based inpainting. In IEEE CVPR, volume 2, pages 721–728, 2003. 1

[6] G. H. Golub and C. F. V. Loan. Matrix Computations. Third edition, 1996. 3, 4, 7

[7] R. I. Hartley and A. Zisserman. Multiple View Geometry in Computer Vision. Cambridge University Press, 2, 3, 4, 5, 6, 7

[8] H. V. Henderson and S. R. Searle. The vec-permutation matrix, the vec operator and Kronecker products: A review. J. Linear Multilinear Algebra, 9:271–288, 1981. 3

[9] V. Jain and P. Narayanan. Video Completion for Indoor Scenes. In ICCVGIP, pages 409–420, 2006. 1

[10] K. Kanatani and N. Ohta. Accuracy bounds and optimal computation of homography for image mosaicing applications. In IEEE ICCV, volume 1, pages 73–78, 1999. 2, 4

[11] S. Kang, T. Chan, and S. Soatto. Inpainting from Multiple View, UCLA CAM Report 02-31, 2002. 1

[12] S. Osher, G. Sapiro, L. Vese, and M. Bertalmio. Simultaneous structure and texture image inpainting. In IEEE CVPR, volume 2, pages 707–712, 2003. 1

[13] S. Periaswamy and H. Farid. Elastic registration with partial data. In International Workshop on Biomedical Image Registration, 2003. 7

[14] E. Rosten and T. Drummond. Machine learning for high-speed corner detection. In European Conf. on Computer Vision, volume 1, pages 430–443, May 2006. 1

[15] Y. Wexler, E. Shechtman, and M. Irani. Space-time completion of video. IEEE Trans. PAMI, 29(3):463–476, March 2007. 1

[16] S. Zokai and G. Wolberg. Image registration using log-polar mappings for recovery of large-scale similarity and projective transformations. IEEE Trans. on Image Processing, 14(10):1422–1434, 2005. 1