Single View Modeling

Topics in Image-Based Modeling & Rendering

CSE291 J00

Lecture 6

Jin-Su Kim

This Lecture

- A. Criminisi, I. Reid and A. Zisserman. Single View Metrology, Proc. IEEE International Conference on Computer Vision, 1999
- B.M. Oh, M. Chen, J. Dorsey, F. Durand, Image-Based Modeling and Photo Editing. ACM SIGGRAPH 2001
- Other References

[1] R.T. Collins and R.S. Weiss. Vanishing point calculation as a statistical inference on the unit sphere. In Proc. ICCV pages 400-403, Dec 1990

[2] Stan Birchfield. An introduction to projective geometry.
http://robotics.stanford.edu/~birch/projective/

[3] R. Hartley and A. Zisserman. Multiple view geometry in computer vision. Cambridge, 2000

Overview (1)

- Input : A single image obtained by perspective projection
- Objectives :

(1) Analyze the scene
  - Compute 3D measurements
  - Determine the camera's location

(2) Determine 3D structure of the scene
  - 3D reconstruction of the scene

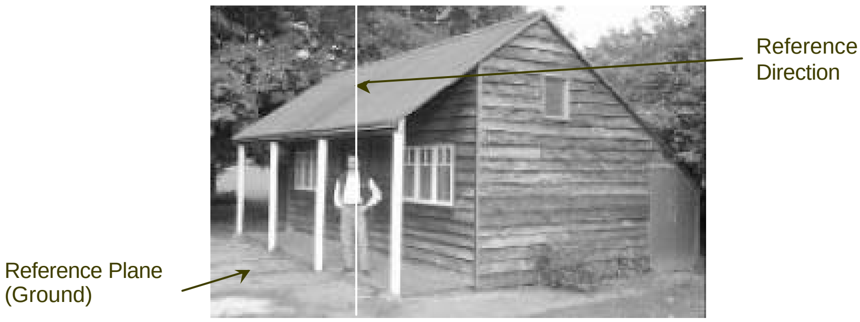
(3) Edit image in 3D
  - Change scene structure, appearance, and illumination

Overview (2)

- Ultimately, build a Photoshop-like tool for image-based modeling and editing.
- Application
  - Forensic science
  - 3D graphical modeling
  - 3D navigation inside a 2D image

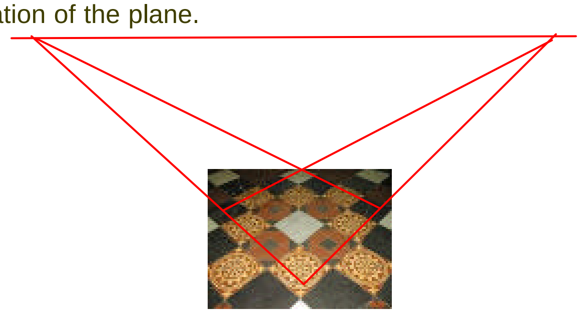
Basic Geometry for Scene Analysis (1)

- Pick a reference plane in the scene.
- Pick a reference direction (not parallel to the reference plane) in the scene.



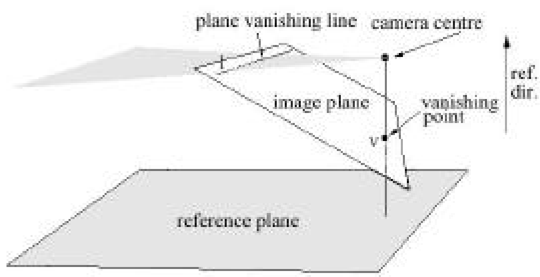
Basic Geometry for Scene Analysis (2)

- Under perspective projection, parallel lines in three-space project to converging lines in the image plane. The common point of intersection, perhaps at infinity, is called the <u>vanishing point</u>. [1]
- Two or more vanishing points from lines known to lie in a single 3D plane establish a <u>vanishing line</u>, which completely determines the orientation of the plane.



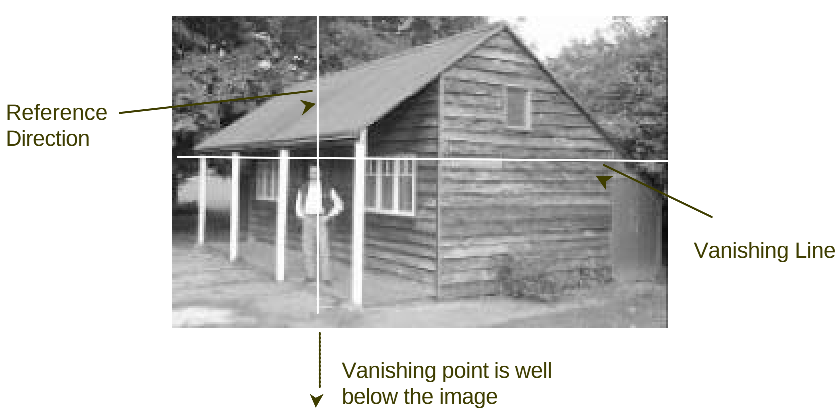
Basic Geometry for Scene Analysis (3)

- The vanishing point for the reference direction is the image of the point at infinity in the reference direction.
- The vanishing line of the reference plane is the projection of the line at infinity of the reference plane into the image.
- The vanishing point and the vanishing line are generally easily obtainable from images of structured scenes.



Basic Geometry for Scene Analysis (4)

- Let's assume that the vanishing point and the vanishing line have been determined and see what can be done with these information.



Three Canonical Types of Measurements

- Measurements of the distance between any of the planes which are parallel to the reference plane
- Measurements on these planes (and comparison of these measurements to those obtained on any plane)
- Determining the camera's position in terms of the reference plane and direction

Measurements between Parallel Planes (1)

- Two points on separate planes (parallel to the reference plane) <u>correspond</u> if the line joining them is parallel to the reference direction.
- For example, if the reference direction is vertical, then the top of an upright person's head and the sole of his/her foot correspond.

Measurements between Parallel Planes (2)

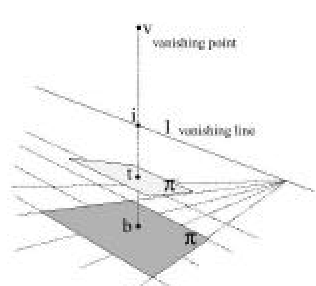
- Projective geometry preserves neither distances nor ratios of distances. However, the <u>cross-ratio</u>, which is a ratio of ratios of distances, is preserved under projective transformations. [2]
- Given four collinear points $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$, and $\mathbf{p}_4$ in P^2, denote the Euclidean distance between two points $\mathbf{p}_i$ and $\mathbf{p}_j$ as $\mathbf{p}_i\mathbf{p}_j$. Then, one definition of the cross-ratio is

$$Cr(\mathbf{p}_1, \mathbf{p}_2; \mathbf{p}_3, \mathbf{p}_4) = (\mathbf{p}_1\mathbf{p}_3 / \mathbf{p}_1\mathbf{p}_4) \;/\; (\mathbf{p}_2\mathbf{p}_3 / \mathbf{p}_2\mathbf{p}_4)$$
$$= (\mathbf{p}_1\mathbf{p}_3 \times \mathbf{p}_2\mathbf{p}_4) \;/\; (\mathbf{p}_1\mathbf{p}_4 \times \mathbf{p}_2\mathbf{p}_3).$$

Measurements between Parallel Planes (3)

- We wish to measure the distance between two parallel planes, specified by the image points $\mathbf{t}$ and $\mathbf{b}$, in the reference direction. (points $\mathbf{t}$ and $\mathbf{b}$ are in correspondence)
- The four points $\mathbf{v}$, $\mathbf{i}$, $\mathbf{t}$, and $\mathbf{b}$ on the image plane define a cross-ratio,

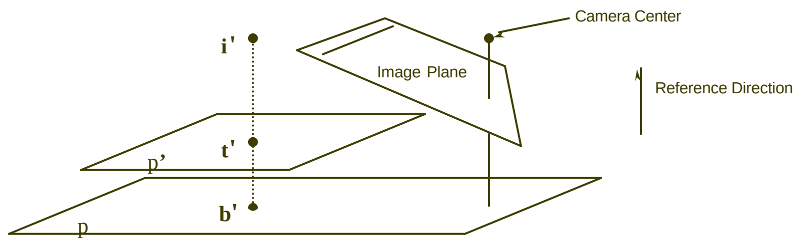
$$Cr(\mathbf{v}, \mathbf{b}; \mathbf{i}, \mathbf{t}) = (\mathbf{vi} / \mathbf{vt}) / (\mathbf{bi} / \mathbf{bt}) = (\mathbf{vi} / \mathbf{vt}) \times (\mathbf{bt} / \mathbf{bi}) .$$



Measurements between Parallel Planes (4)

- Assume that the points **v**, **i**, **t**, and **b** on the image plane are the images of points **v'**, **i'**, **t'**, and **b'** in the world coordinate under projective transformation. Then the cross-ratio of the points in the world coordinate is

$$Cr(\mathbf{v'}, \mathbf{b'}; \mathbf{i'}, \mathbf{t'}) = (\mathbf{v'i'} / \mathbf{v't'}) / (\mathbf{b'i'} / \mathbf{b't'})$$
$$= (\mathbf{v'i'} / \mathbf{v't'}) \times (\mathbf{b't'} / \mathbf{b'i'}).$$

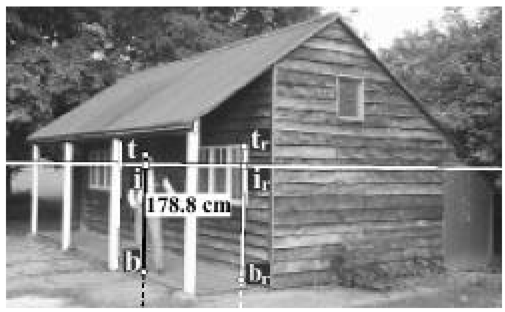


Measurements between Parallel Planes (5)

- Since **v'** is the point at infinity, **v'i'** and **v't'** cancel out each other. Also note that **b'i'** is the camera's distance from the plane p, and **b't'** is the distance between the planes p and p'.
- Since the cross-ratio is invariant under projective transformation, we can obtain the absolute distance between the planes once the camera's distance from p is specified.
- It is more practical to determine the distance via a second measurement in the image, that of a known reference length.

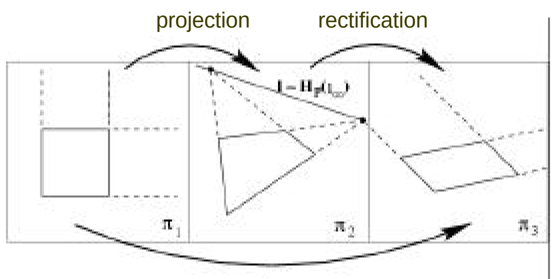
Measurements between Parallel Planes (6)

- A person's height can be computed from an image given a vertical reference height elsewhere in the scene.
- The reference height is the segment ($\mathbf{t}_r$, $\mathbf{b}_r$).
- $\mathbf{t}$ is the top of the head and $\mathbf{b}$ is the base of the feet of the person while $\mathbf{i}$ is the intersection with the vanishing line.



Measurements on Parallel Planes (1)

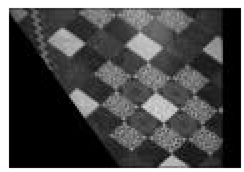
- A projective transformation maps $\mathbf{l}_\infty$ (the line at infinity) from $(0, 0, 1)^T$ on a Euclidean plane π_1 to a finite line $\mathbf{l}$ on the plane π_2.
- If a projective transformation is constructed such that $\mathbf{l}$ is mapped back to $(0, 0, 1)^T$ then the transformation between the first and third planes is an affine transformation. This is called the affine rectification. [3]



Measurements on Parallel Planes (2)

- If we know the vanishing line of the reference plane then by the affine rectification, we can measure
  - ratios of lengths of parallel line segments on the plane
  - ratio of areas on the plane
- Moreover, since the vanishing line is shared by the pencil of planes parallel to the reference plane, these ratios may be obtained for any other plane in the pencil.
- How about between two parallel planes?

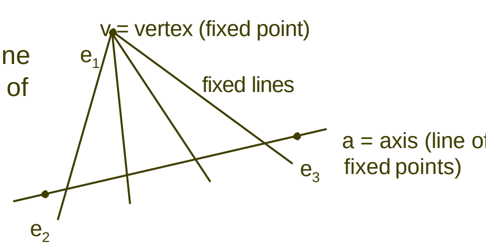
Projected Image



Rectified Image

Measurements on Parallel Planes (3)

- A plane projective transformation is a planar homology if it has a line of fixed points (called the axis), together with a fixed point (called the vertex) not on the line. [3]
- There is a pencil of fixed lines through the vertex.
- Algebraically, two of the eigenvalues of the transformation matrix are equal, and the fixed line corresponds to the 2D invariant space of the matrix. The vertex corresponds to the other eigenvalue.
- The ratio of the distinct eigenvalue to the repeated one is the characteristic invariant of the homology.

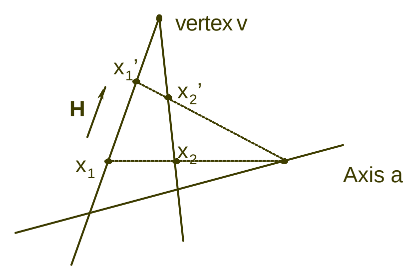


Measurements on Parallel Planes (4)

- Properties of a planar homology :

1) Lines joining corresponding points intersect at the vertex, and corresponding lines intersect on the axis.

- Under the transformation points on the axis are mapped to themselves.
- Each point off the axis lies on a fixed line through **v** intersecting **a** and is mapped to another point on the line.
- Consequently, corresponding point pairs $\mathbf{x} \leftrightarrow \mathbf{x}'$ and the vertex of the homology are collinear.
- Corresponding lines (i.e. lines through pairs of corresponding points) intersect on the axis: for example, the lines <$\mathbf{x}_1$, $\mathbf{x}_2$> and <$\mathbf{x}_1$', $\mathbf{x}_2$'>.

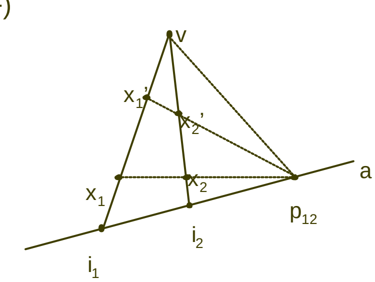


Measurements on Parallel Planes (5)

- Properties of a planar homology :

2) The cross ratio defined by the vertex v, the corresponding points **x, x'**, and the intersection of their join with the axis **i**, is the characteristic invariant of the homology, and is the same for all points related by the homology.

3) The vertex (2 DOF), axis (2 DOF) and invariant (1 DOF) are sufficient to define the homology completely. A planar homology thus has 5 degrees of freedom.

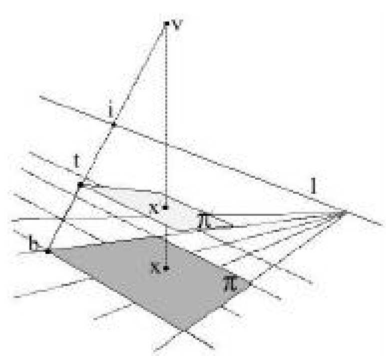


Measurements on Parallel Planes (6)

- A map in the world between parallel planes induces a map between images of the two planes. This image map is a planar homology.
- In our case the vanishing line of the plane (2 DOF), and the vertical vanishing point (2 DOF), are respectively, the axis and vertex of the homology which relates a pair of planes in the pencil.
- The remaining 1 DOF of the homology is uniquely determined from any pair of image points which correspond between the planes.

Measurements on Parallel Planes (7)

- In the figure, the mapping between the images of the two planes is a homology, with **v** the vertex and **l** the axis.
- The correspondence $\mathbf{b} \rightarrow \mathbf{t}$ fixes the remaining degree of freedom of the homology from the cross-ratio of the four points: **v**, **i**, **t** and **b**.



Measurements on Parallel Planes (8)

- We can compare measurements made on two separate planes by mapping between the planes in the reference direction via the homology.
- Simply transfer all points from one plane to the reference plane using the homology, then use the previous method.
- In particular, we may compute
  - the ratio between two parallel lengths, one length on each plane
  - the ratio between two areas, one area on each plane.

Determining the Camera Position

- Conversing the procedure of “Measurements between parallel planes”, we may compute the camera’s distance from the reference plane knowing a single reference distance.
- The location of the camera relative to the reference plane is the back-projection of the vanishing point onto the reference plane.
- This back-projection is accomplished by a homography which maps the image to the reference plane.

Application : Modeling Paintings

- *La Flagellazione di Cristo* (1460) Galleria Nazionale delle Marche by Piero della Francesca (1416-1492)

- *La Trinita'* (1426) Firenze, Santa Maria Novella; by Masaccio (1401-1428)

- Visit http://www.robots.ox.ac.uk/~vgg/projects/SingleView for more examples.

Image-Based Modeling and Editing System (1)

- An interactive modeling and editing system that uses an image-based representation for the entire 3D authoring process
- Input : A single photograph
- Features

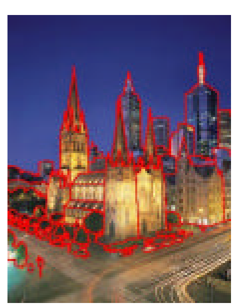
  - Various editing operations such as painting, copy-pasting, and relighting.
  - Tools to extract layers and assign depths
  - Editing from different viewpoints
  - Extracting and grouping of image-based objects
  - Modifying the shape, color, and illumination of the objects

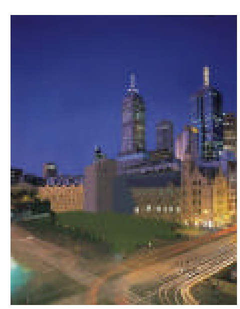
Image-Based
Modeling and Editing System (2)

- Non-distorted clone brushing
  - Permits the distortion-free copying of parts of a picture
  - By using parameterization optimization technique.
- Texture-illuminance decoupling filter
  - Discounts the effect of illumination on uniformly texture areas
  - By decoupling large- and small-scale features via bilateral filtering.

System Overview (1)

- All elements of the system operate on the same data structure : images with depth.
- A scene is represented as a layered collection of depth images by
  1) Manually segmenting the original image into different layers. Layering is done at a high level (objects or object parts)
  2) Manually painting the parts of the scene hidden in the input image using clone brushing
  3) Assigning depths to objects for each layer.





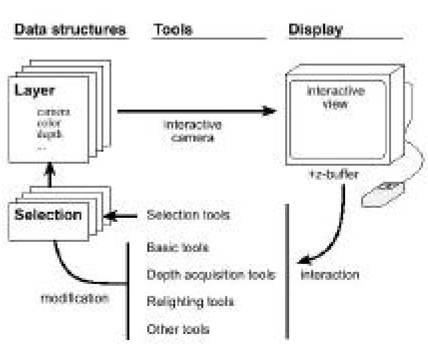
System Overview (2)

- Each pixel of a layer encodes color, depth, etc.
- Alpha channel is used to handle transparency and object masks.
- Each layer has a reference camera that describes its world-to-image projection matrix. Change of this matrix causes changes of channel values.
  - Reference images correspond to the original matrix.
  - Interactive images are displayed from different viewpoints to ease user interaction

```
layer {
    reference camera  : transformation matrix
    color channels    : array of floats
    alpha channel     : array of floats
    depth channel     : array of floats
    optional channels : arrays of floats
}
```

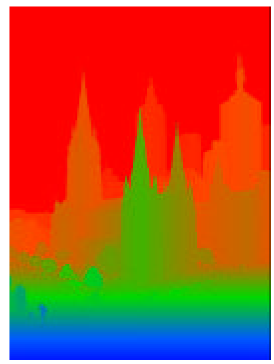
System Overview (3)

- The system consists of a set of tools organized around a common data structure.
- The tools, such as depth assignment, selection or painting can be used from any interactive viewpoint.



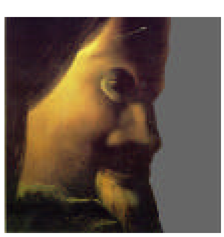
Depth Assignment (1)

- Paint depth like colors are painted.
- Relies on the user's ability to comprehend the layout of the scene.
- The level of detail and accuracy of depth depend on the target application.
- Use cues present in the image directly. Use previously-assigned depth as a reference.
- Depth can be edited from any interactive viewpoint. Multiple views can also be used.



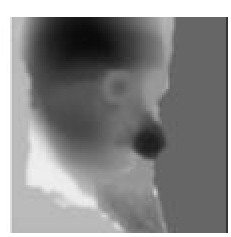
Depth Assignment (2)

- The system provides hybrid tools that aid in assigning accurate depth.
- The user can directly paint depth using a brush, either
  - setting the absolute depth value or
  - adding or subtracting to the current value (<u>chiseling</u>).
- <u>Blurring</u> smoothes the shape, while <u>sharpening</u> accentuates relief.



Face

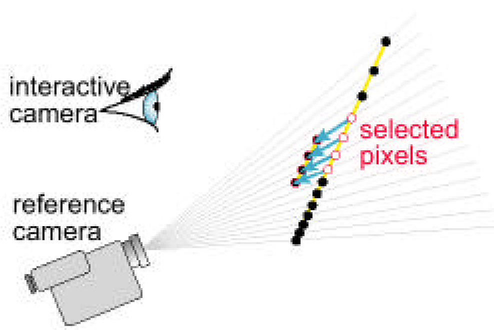
Chiseled depth



Blurred depth

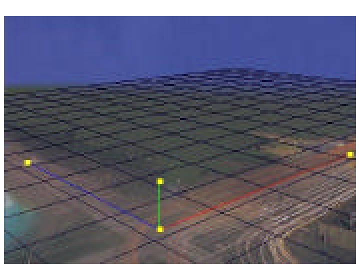
Depth Assignment (3)

- The whole selected region can be translated along lines of sight with respect to the reference camera. (depth translation)
- Depth translation is done by multiplying the depth of each selected pixel by a constant value. Hence depth-translating planar objects results in parallel planar objects.



Depth Assignment (4)

- The use of a reference ground plane greatly simplifies depth acquisition and improves accuracy dramatically.
- Specifying a ground plane is typically the first step of depth assignment.

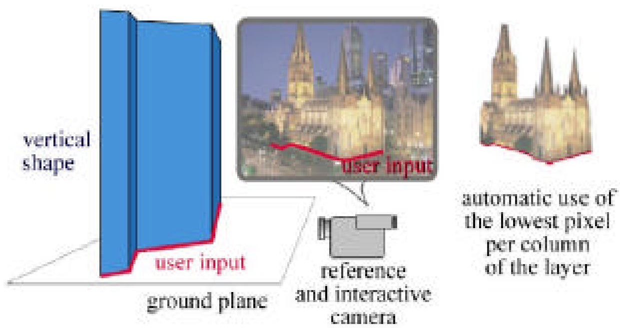


Ground Plane

Depth Map

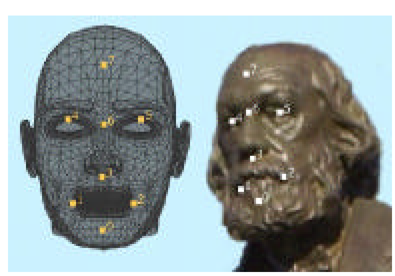
Depth Assignment (5)

- The user draws the contact of the vertical geometry with the ground plane.



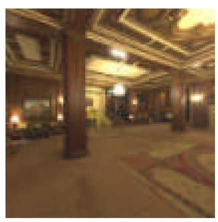
Depth Assignment (6)

- Geometric primitives
  - To depth-paint geometric shapes such as boxes, spheres, or cylinders, the user draws 2D geometric primitives.
- Organic shapes
  - Use level sets to give a bulgy appearance by specifying more distant depth at the boundary and closer depth in the center.
- Faces
  - Use a generic arbitrary 3D face model.
  - The user specifies correspondence points between the image and the 3D model.

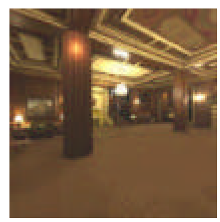


Non-distorted Clone Brushing (1)

- Limitations of standard clone brushing
  - Only parts with similar orientation and distance to the camera can be copied because perspective causes the texture foreshortening.
  - Only regions of similar intensity can be copied.
- Objective : Map the source region of the image-based representation to the destination, with as little distortion as possible.



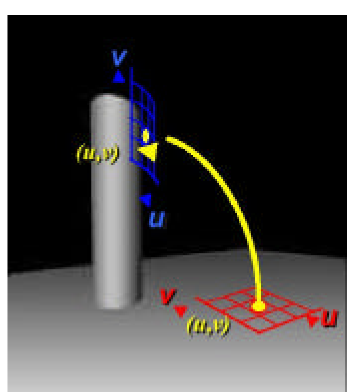
Initial Image



Clone-Brushed Image

Non-distorted Clone Brushing (2)

- Idea : Compute a (u, v) texture parameterization for both the source and destination regions, and use this mapping for the clone brush.



Non-distorted Clone Brushing (3)

- Flood-fill parameterization

  - Adapt the discrete smooth interpolation method (by Levy et al.) in a "flood-fill" manner to optimize the parameterization around the current position of the clone brush.
  - Compute the parameterization for only a subset of pixels, called active pixels.
  - Optimization step : coordinates are refined.
  - Expansion step : new pixels are declared active and initialized.
  - Freezing : freeze the coordinates of already brushed pixels.

Texture-illuminance Decoupling Filter (1)

- An image-processing filter factors the image into a texture component and an illumination component.
- Remove lighting effects from uniformly texture objects.
- Useful both for relighting and clone brushing
- Assumption

  - Large-scale luminance variations are due to the lighting, while small-scale details are due to the texture.
  - The average color comes from the texture component.
- Limitation : Small detailed shadows are not handled correctly.

Texture-illuminance Decoupling Filter (2)

- Procedure :

1) The user specifies a feature size of the texture by dragging a line segment over a pattern.

2) Blur the image with a low-pass Gaussian filter specified by the feature size. => Only large-scale illuminance variations remain.

3) Divide the illuminance obtained by the normalized average color value.

4) Divide the initial image by this blurred version to compute a uniform texture component.

Texture-illuminance Decoupling Filter (3)

- Problem :

- Texture foreshortening needs to be treated to make consistent use of the feature size.

- Shadow boundaries introduce frequencies that are in the range of the feature size. => treated as texture frequencies

- Depth Correction : To compensate for foreshortening

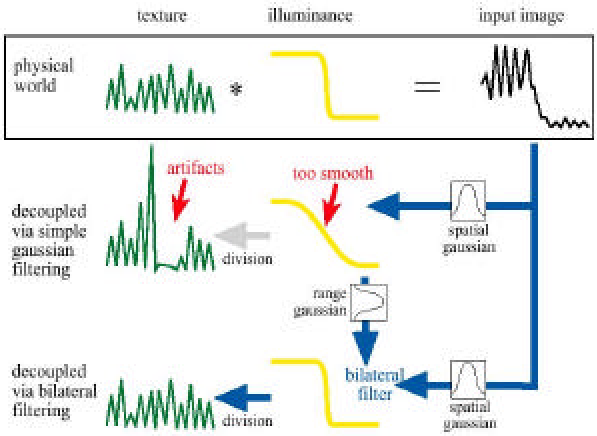
- The user specifies feature size at a reference pixel p_{ref}.

- For other pixels, the spatial kernel is scaled by $\mathbf{z}_{ref} / \mathbf{z}$.

- Bilateral filtering

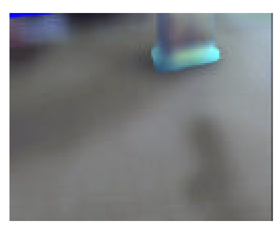
- To handle discontinuities, use a non-linear edge-preserving filter.

Texture-illuminance Decoupling Filter (4)



Texture-illuminance Decoupling Filter (5)

Input Image



Illumination after simple Gaussian filtering

Texture after simple Gaussian filtering

Texture after bilateral filtering

Demo

- Movie Clip
- Visit http://graphics.lcs.mit.edu/ibedit/