The Quotient Image: Class-based Recognition and Synthesis Under Varying Illumination

T. Riklin-Raviv and A. Shashua

Institute of Computer Science
Hebrew University
Example: 3 basis images

The movie is created by simply linearly combining the 3 basis images.
Image Synthesis:

Given a single image $y =$

And a database of other images of the same “class”

We would like to generate new images from $y$ simulating change of illumination.
Recognition

Given a database of images of the same “class”, under varying illumination conditions, and a novel image.

Match between images of the same object.
Definition: Ideal Class of Objects

The images produced by an ideal class of objects are

\[ I(x, y) = \rho_i(x, y)n(x, y)^T s_j \]

Where \( \rho_i(x, y) \) is the albedo (surface texture) of object \( i \) of the class

\( n(x, y) \) is the normal direction (shape) of the surface, i.e., all objects have the same shape.

\( s_j \) is the point light source direction.
Related Work

- Basic result: Shashua 91, 97
- Application and related systems: Hallinan 94, Kriegman et. al 96, 98
- Rendering under more general assumption: Dorsey et al. 93, 94
- Work on “class-based” synthesis and recognition of images - mostly with varying viewing positions: Basri 96, Bymer & Poggio 95, Freeman & Tenenbaum 97, Vetter & Blanz 98, Vetter, Jones & Poggio 97, 98, Edelman 95, Atick, Griffin & Redlich 97.
- Linear class: Vetter & Poggio 97, 92
- Additive error term: Sali & Ulman 98
- Reflectance Ratio: Nayar & Bolle 96
Basic Assumptions

- Lambertian surface, linear model: no cast shadows, no highlights.
- Same view point (canonical view).
- Ideal class assumption
- Images have the same size and are (roughly) aligned.
The Quotient image: Definition

Given images \( y_s, a_s \)

of objects \( y \) and \( a \) respectively, under illumination \( S \)

\[
\begin{align*}
    y_s &= \rho_y N^T s \\
    a_s &= \rho_a N^T s
\end{align*}
\]

Thus \( Q_y \) depends only on relative surface information and is independent of illumination.
The Quotient image Method: Proposition

Let $a_1, a_2, a_3$ 3 images of object $a$.

Let $y_s$ Image of object $y$ illuminated by light $s$.

Then, there exist $x_1, x_2, x_3$ that satisfy:

$$y_s = (\sum_j x_j a_j) \otimes Q_y$$

Moreover, the image space of $y$ is spanned by varying the coefficients.
The Quotient image Method: Proof

\(a_1, a_2, a_3\)  Illuminated by: \(s_1, s_2, s_3\)

\(y_s\)  Illuminated by: \(s = \sum_j x_j s_j\)

\[y_s = (\sum_j x_j a_j) \otimes Q_y\]

\[\rho_a n^T s \rho_y / \rho_a\]
The Quotient image Method: \( N=1 \)

\[ s = \sum_j x_j s_j \]

\[ a_s = x_1 a_1 + x_2 a_2 + x_3 a_3 \]
The Quotient image Method: \( N=1 \)

\[ a_{s'} = z_1 a_1 + z_2 a_2 + z_3 a_3 \]

\[ a_{s'} \quad \text{Q-Image} \quad \text{Synthesized Image} \]
The Quotient image Method: Conclusions

- Given $Q_y$, one can generate $Y_s$ and all other images of the image space of $y$.

- Given $Y_s$ and the coefficients $x_j$ that satisfies

$$s = \sum_j x_j s_j \quad \text{then} \quad Q_y \quad \text{readily follows} \quad Q_y = \frac{y_s}{\sum_j x_j a_j}$$

- In order to obtain the correct coefficients $x_j$

a bootstrap set of more than one object is needed.
The Quotient image Method: $N>1$

Original image

$$f(\hat{x}) = \frac{1}{2} \sum_{i=1}^{N} |A_i \hat{x} - \alpha_i y_s|^2$$

$A_i$ is $m \times 3$
The Quotient image Method:
Theorem-1

The energy function

$$f(\hat{x}) = \frac{1}{2} \sum_{i=1}^{N} |A_i \hat{x} - \alpha_i y_s|^2$$

has a (global) minimum \( x = \hat{x} \), if the albedo \( \rho_y \) of object \( y \) is rationally spanned by the bootstrap set.

i.e if there exist coefficients \( \alpha_1, \ldots, \alpha_N \) such that

$$\rho_y = \frac{\rho_1^2 + \ldots + \rho_N^2}{\alpha_1 \rho_1 + \ldots + \alpha_N \rho_N}$$
The Quotient Image Method: 
Solving For $X$ and $\alpha_i$

$$\min_{x,\alpha} f(\hat{x}) = \frac{1}{2} \sum_{i=1}^{N} |A_i \hat{x} - \alpha_i y_s|^2$$

$$\hat{x} = \left( \sum_i A_i^T A_i \right)^{-1} \left( \sum_i \alpha_i A_i^T \right) y_s = \sum_i \alpha_i v_i$$

$$v_i = \left( \sum_{r=1}^{N} A_r^T A_r \right)^{-1} A_i^T y_s$$

We also have: $$0 = \frac{\partial f}{\partial \alpha_i} = \alpha_i y_s^T y_s - \hat{x}^T A_i^T y_s$$
The Quotient image Method: Solving For $X$ and $\alpha_i$

for $i = 1 \ldots N$ written explicitly

$$\alpha_1 \left( v_1^T A_1^T y_s - y_s^T y_s \right) + \ldots + \alpha_N v_N^T A_1^T y_s = 0$$

$$\alpha_1 v_1^T A_2^T y_s + \ldots \ldots \ldots + \alpha_N v_N^T A_2^T y_s = 0$$

$$\vdots$$

$$\alpha_1 v_1^T A_N^T y_s + \ldots + \alpha_N \left( v_N^T A_N^T y_s - y_s^T y_s \right) = 0$$
The Algorithm

Given: $A_1, \ldots, A_N$ a bootstrap set and a novel image $y_s$

Use the minimization function:

$$\min_{x, \alpha} f(\hat{x}) = \frac{1}{2} \sum_{i=1}^{N} |A_i \hat{x} - \alpha_i y_s|^2$$

to generate a homogenous system of linear equations in $\alpha_1, \ldots, \alpha_N$

Scale such that $\sum_i \alpha_i = N$

Compute $x = \sum_i \alpha_i v_i$

$Q_y = \frac{y_s}{Ax}$ Where $A$ is the average of $A_1, \ldots, A_N$

$y_{new}(Z) = Az \otimes Q_y$ For all choices of $z$
Frontal faces: Collection of objects all have the same shape but differ in their surface texture (albedo)...

Samples of few faces out of 9*200 faces images taken from T. Vetter database which was mainly used as a bootstrap set and as a source for novel images in the further demonstration.
Synthesis from Single Picture
And 10 faces from the bootstrap set
under 3 different light conditions

Synthesis from 3 pictures
10 other faces from the database, each under 3 light conditions.

Synthesis from Single image and the bootstrap set

Synthesis from 3 pictures
Synthesis from 3 pictures

Synthesis from Single image and the bootstrap set
Original Images Compared to Q-Image Synthesized Images

1st Row: Original Images
2nd Row: Q-Image Synthesized Images
3rd Row: Exact Values of Light Direction: center, down, up, right, left
Light Coefficient Comparison Ground Truth Vs. Q-Image Coefficients

1st Coefficient

2nd Coefficient

3rd Coefficient
Using Different Database

The Quotient Image

The original image

3x3 images’ Database

Animation Using the database
Handling Color Images

RGB  ↔  HSV  Transformation

Original color image  R  G  B

H  S  V
Original Image

Quotient Image

Synthesized Sequence
Monica and Bill Under a New Light

Original Images

Quotient Images

Synthesized Sequences
Recognition under varying illumination

Database generation

Each object in the database is represented by its quotient image only. The quotients can be made of images with varying illuminations.

The quotient images were generated out of $N \times 3$ ($N=20$) base images.
**Identification**

Given a new image of an object appears in the data base under any light condition, it’s quotient is computed from A,B,C … (as was done in the database generation) . Then It is compared to the quotients in the data base.

**Other methods used for comparison**

1. **Correlation**
   
   **Database:** Each object is represented in the database by it’s image under any/certain lightening condition.
   
   **Identification:** Correlation between the test image to the images stored in the database.

2. **PCA**
   
   **Database:** Applying PCA on the objects’ images + 3*20 additional images of 20 objects under 3 illumination (to compare conditions to the quotient method). Having eigen vectors, each object is represented by it’s eigen vectors’ coefficients.
   
   **Identification:** Comparison (LSE) between the test image coefficients (generated the same way as the database) and the database.
Recognition Results

Quotient method comparing to correlation
Recognition Results - cont

Quotient Method Vs. PCA
The End