

# Image and Depth from a Conventional Camera with a Coded Aperture

Pramook Khungurn

December 11, 2019

This is an “untangled” version of “Image and Depth from a Conventional Camera with a Coded Aperture” by Leven et al.[3]. In this note, I try to explain some concepts that are new to myself at the time of reading.

## 1 Statistical Model of Images

- The paper thinks of an image as an array of real numbers. Of course, the array is 2D, but it’s easier to think of it as a 1D vector  $x = (x[1, 1], x[1, 2], \dots, x[p, q])^T \in \mathbb{R}^{pq}$ .
- Citing a psychology paper from Cornell [4], the paper claims that natural images have sparse derivative distribution. However, to make computation tractable, they assume that the derivative has Gaussian distribution.

That is, the property we observe a particular pixel to be a particular value is given by:

$$\Pr(x[i, j]) \propto \exp \left( -\frac{1}{2} (\alpha(x[i, j] - x[i + 1, j])^2 + (x[i, j] - x[i, j + 1])^2) \right),$$

where  $\alpha$  is set to the variance of distribution that matches the variance of derivatives of natural images. (The paper sets  $\alpha = 250$ .)

Therefore, the probability we observe the whole image is given by:

$$\Pr(x) = \prod_{i,j} \Pr(x[i, j]) = \prod_{i,j} \exp \left( -\frac{1}{2} (\alpha(x[i, j] - x[i + 1, j])^2 + (x[i, j] - x[i, j + 1])^2) \right)$$

- Let  $g_x$  be vector corresponding to the filter that computes the derivative along  $x$ -direction. That is,

$$g_x[i, j] = \begin{cases} 1, & \text{if } i = j = 0 \\ -1, & \text{if } i = 1 \text{ and } j = 0. \end{cases}$$

Let  $C_{g_x}$  be the  $pq \times pq$  matrix of the convolution with  $g_x$ . That is,

$$C_{g_x}[(i, j), (i', j')] = \begin{cases} 1, & \text{if } (i, j) = (i', j') \\ -1, & \text{if } (i + 1, j) = (i', j') \\ 0, & \text{otherwise.} \end{cases}$$

We also define  $g_y$  and  $C_{g_y}$  in a similar manner for derivative along the  $y$  direction.

- Now, we have that

$$\begin{aligned}
\Pr(x) &\propto \prod_{i,j} \exp \left( -\frac{1}{2} (\alpha(x_{i,j} - x_{i+1,j})^2 + (x_{i,j} - x_{i,j+1})^2) \right) \\
&= \exp \left( -\frac{\alpha}{2} \sum_{i,j} ((x_{i,j} - x_{i+1,j})^2 + (x_{i,j} - x_{i,j+1})^2) \right) \\
&= \exp \left( -\frac{\alpha}{2} \left( \sum_{i,j} (x_{i,j} - x_{i+1,j})^2 + \sum_{i,j} (x_{i,j} - x_{i,j+1})^2 \right) \right) \\
&= \exp \left( -\frac{\alpha}{2} \left( x^* C_{g_x}^* C_{g_x} x + x^* C_{g_y}^* C_{g_y} x \right) \right) \\
&= \exp \left( -\frac{1}{2} \left( x^* (\alpha(C_{g_x}^* C_{g_x} + C_{g_y}^* C_{g_y})) x \right) \right) \\
&= \exp \left( -\frac{1}{2} x^* \Psi^{-1} x \right)
\end{aligned}$$

where  $\Psi^{-1} = \alpha(C_{g_x}^* C_{g_x} + C_{g_y}^* C_{g_y})$ . That is,  $\Pr(x) \propto N(\mathbf{0}, \Psi)$  where  $N$  is the anisotropic Gaussian distribution.

- We now would like to describe the probability distribution of the Fourier transform  $X = \mathcal{F}x$ . Note that  $X$  is still an  $pq$ -vector. However, the elements of  $X$  are denoted by  $(\nu, \omega)$  instead of  $(i, j)$ . Note also that we write  $\mathcal{F}x$  instead of the standard notation  $\mathcal{F}\{x\}$ . This is because the Fourier transform is a linear transformation and thus can be represent by a matrix. So,  $\mathcal{F}$  is an  $pq \times pq$  matrix in this case.
- Using this notation, we can state the Fourier convolution theorem in another way. Let  $g$  be a convolution filter, and let  $C_g$  be its corresponding matrix. Then, we have that

$$\mathcal{F}C_g = \text{diag}(G)\mathcal{F}$$

for  $G = \mathcal{F}g$ . In other words,  $C_g = \mathcal{F}^{-1}\text{diag}(G)\mathcal{F}$ , which says that  $C_g$  has an eigen decomposition. (This is not a surprise because  $C_g$  is Toeplitz, and Toeplitz matrices have Fourier vectors as eigenvalues.)

- Because  $C_{g_x}$  represents a convolution, we have that  $\mathcal{F}C_{g_x}x = \text{diag}(G_x)X$  where  $G_x$  is the Fourier transform of  $g_x$ . We can also define  $G_y$  in an analogous manner.
- Let  $c$  be the constaint such  $\mathcal{F}^*\mathcal{F} = cI$ . The probability distribution of the Fourier transform is given by:

$$\begin{aligned}
\Pr(X) &= \Pr(x) \propto \exp \left( -\frac{\alpha}{2} x^* (C_{g_x}^* C_{g_x} + C_{g_y}^* C_{g_y}) x \right) = \exp \left( -\frac{\alpha}{2c} x^* (C_{g_x}^* (cI) C_{g_x} + C_{g_y}^* (cI) C_{g_y}) x \right) \\
&= \exp \left( -\frac{\alpha}{2c} x^* (C_{g_x}^* \mathcal{F}^* \mathcal{F} C_{g_x} + C_{g_y}^* \mathcal{F}^* \mathcal{F} C_{g_y}) x \right) \\
&= \exp \left( -\frac{\alpha}{2c} x^* (\mathcal{F}^* \text{diag}(G_x)^* \text{diag}(G_x) \mathcal{F} + \mathcal{F}^* \text{diag}(G_y)^* \text{diag}(G_y) \mathcal{F}) x \right) \\
&= \exp \left( -\frac{\alpha}{2c} x^* \mathcal{F}^* (\text{diag}(G_x)^* \text{diag}(G_x) + \text{diag}(G_y)^* \text{diag}(G_y)) \mathcal{F} x \right) \\
&= \exp \left( -\frac{\alpha}{2c} X^* (\text{diag}(G_x)^* \text{diag}(G_x) + \text{diag}(G_y)^* \text{diag}(G_y)) X \right) \\
&= \exp \left( -\frac{1}{2} X^* \bar{\Psi}^{-1} X \right)
\end{aligned}$$

where  $\bar{\Psi}^{-1} = \frac{\alpha}{c}(\text{diag}(G_x)^*\text{diag}(G_x) + \text{diag}(G_y)^*\text{diag}(G_y))$ . So, the Fourier transform has a Gaussian distribution as well.

Also, from the above derivation, notice that  $x^*\Psi^{-1}x = X^*\bar{\Psi}^{-1}X = x^*\mathcal{F}^*\bar{\Psi}^{-1}\mathcal{F}x$ . In other words,

$$\begin{aligned}\Psi^{-1} &= \mathcal{F}^*\bar{\Psi}^{-1}\mathcal{F} \\ \mathcal{F}^{-*}\Psi^{-1}\mathcal{F}^{-1} &= \bar{\Psi}^{-1} \\ \mathcal{F}\Psi\mathcal{F}^* &= \bar{\Psi}.\end{aligned}$$

- Suppose we have a blurred kernel  $f_k$ . We will observe the signal

$$y = f_k * x + n = C_{f_k}x + n$$

where  $n$  is a noise with Gaussian distribution  $n \sim N(\mathbf{0}, \eta^2 I)$ . The noise is assumed to be independent of the other part of the signal.

- Since  $C_{f_k}x$  is a linear transformation of  $x$  and  $x$  has Gaussian distribution  $N(\mathbf{0}, \Psi)$ . We have that  $Pr(C_{f_k}x) \sim N(\mathbf{0}, C_{f_k}\Psi C_{f_k}^*)$ . (This is a theorem?)

Now,  $y = C_{f_k}x + n$ . Since  $n$  is indenpent from  $C_{f_k}x$ , we have that  $Pr(y) \sim N(\mathbf{0}, C_{f_k}\Psi C_{f_k}^* + \eta^2 I)$ . We let  $\Sigma_k = C_{f_k}\Psi C_{f_k}^* + \eta^2 I$ .

- Let  $Y = \mathcal{F}y$ . We have that

$$\begin{aligned}\Pr(Y) &= \Pr(y) \propto \exp\left(-\frac{1}{2}y^*(C_{f_k}\Psi C_{f_k}^* + \eta^2 I)^{-1}y\right) \\ &= \exp\left(-\frac{1}{2}y^*\mathcal{F}^*\mathcal{F}^{-*}(C_{f_k}\Psi C_{f_k}^* + \eta^2 I)^{-1}\mathcal{F}^{-1}\mathcal{F}y\right) \\ &= \exp\left(-\frac{1}{2}Y^*(\mathcal{F}(C_{f_k}\Psi C_{f_k}^* + \eta^2 I)\mathcal{F}^*)^{-1}Y\right) \\ &= \exp\left(-\frac{1}{2}Y^*(\mathcal{F}C_{f_k}\Psi C_{f_k}^*\mathcal{F}^* + \eta^2\mathcal{F}\mathcal{F}^*I)^{-1}Y\right) \\ &= \exp\left(-\frac{1}{2}Y^*(\mathcal{F}C_{f_k}\Psi C_{f_k}^*\mathcal{F}^* + \eta^2\mathcal{F}\mathcal{F}^*I)^{-1}Y\right) \\ &= \exp\left(-\frac{1}{2}Y^*(\text{diag}(F_k)\mathcal{F}\Psi\mathcal{F}^*\text{diag}(F_k)^* + c\eta^2 I)^{-1}Y\right) \\ &= \exp\left(-\frac{1}{2}Y^*(\text{diag}(F_k)\bar{\Psi}\text{diag}(F_k)^* + c\eta^2 I)^{-1}Y\right).\end{aligned}$$

So,  $Y \sim N(\mathbf{0}, \text{diag}(F_k)\bar{\Psi}\text{diag}(F_k)^* + c\eta^2 I)$ . Let  $\bar{\Sigma}_k = \text{diag}(F_k)\bar{\Psi}\text{diag}(F_k)^* + c\eta^2 I$ .

- Note that, because  $\bar{\Psi}$  is diagonal, we have that  $\bar{\Sigma}_k$  is diagonal as well. Because

$$\bar{\Psi}^{-1} = \frac{\alpha}{c}(\text{diag}(G_x)^*\text{diag}(G_x) + \text{diag}(G_y)^*\text{diag}(G_y)),$$

we have that

$$\bar{\Psi}^{-1}[(\nu, \omega), (\nu, \omega)] = \frac{\alpha}{c}(|G_x[\nu, \omega]|^2 + |G_y[\nu, \omega]|^2).$$

So,

$$\bar{\Psi}[(\nu, \omega), (\nu, \omega)] = \frac{c}{\alpha|G_x[\nu, \omega]|^2 + \alpha|G_y[\nu, \omega]|^2},$$

and

$$\bar{\Sigma}_k[(\nu, \omega), (\nu, \omega)] = \frac{c|F_k[\nu, \omega]|^2}{\alpha|G_x[\nu, \omega]|^2 + \alpha|G_y[\nu, \omega]|^2} + c\eta^2.$$

For brevity, we let  $\sigma_k[\nu, \omega]$  denote  $\bar{\Sigma}_k[(\nu, \omega), (\nu, \omega)]$ .

- Using the notation above, we have that

$$\Pr(Y) = \exp\left(-\frac{1}{2}Y^*\bar{\Sigma}_k^{-1}Y\right) = \exp\left(-\frac{1}{2}\sum_{\nu, \omega}\frac{|Y[\nu, \omega]|^2}{\sigma_k[\nu, \omega]}\right)$$

For brevity, let  $P_k(Y)$  denote the above distribution for a specific filter  $f_k$ .

## 2 Filter Design

- Suppose we have a collection of filter  $f_{k_1}, f_{k_2}, \dots, f_{k_\ell}$ . In the paper, these filters correspond to the impulse responses of the aperture taken at different distances from the target. In other words, they are scaled versions of one another. We would like these filters to be as different as possible.
- Given two filters  $f_{k_i}$  and  $f_{k_j}$ , we measure how different they are by the *Kullback-Leibler divergence* of their corresponding probability distributions:

$$D_{KL}(P_{k_i}, P_{k_j}) = \int_Y P_{k_i}(Y)(\log P_{k_i}(Y) - \log P_{k_j}(Y)) dY.$$

According to the last section,  $P_{k_i}$  and  $P_{k_j}$  are anisotropic Gaussian distributions. Wikipedia says that the KL divergence of such distributions are:

$$D_{KL}(P_{k_i}, P_{k_j}) = \frac{1}{2}\left(\text{tr}(\bar{\Sigma}_{k_j}^{-1}\bar{\Sigma}_{k_i}) + (\mu_{k_j} - \mu_{k_i})^*\bar{\Sigma}_{k_j}^{-1}(\mu_{k_j} - \mu_{k_i}) - \ln\left(\frac{\det \bar{\Sigma}_{k_i}}{\det \bar{\Sigma}_{k_j}}\right) - pq\right).$$

where  $\mu_{k_i}$  and  $\mu_{k_j}$  are the mean of  $P_{k_i}$  and  $P_{k_j}$ , respectively. Because the means are zero,

$$\begin{aligned} D_{KL}(P_{k_i}, P_{k_j}) &= \frac{1}{2}\left(\text{tr}(\bar{\Sigma}_{k_j}^{-1}\bar{\Sigma}_{k_i}) - \ln\left(\frac{\det \bar{\Sigma}_{k_i}}{\det \bar{\Sigma}_{k_j}}\right) - pq\right) \\ &= \frac{1}{2}\left(\sum_{\nu, \omega}\frac{\sigma_{k_i}[\nu, \sigma]}{\sigma_{k_j}[\nu, \sigma]} - \sum_{\nu, \omega}\ln\left(\frac{\sigma_{k_i}[\nu, \sigma]}{\sigma_{k_j}[\nu, \sigma]}\right) - pq\right). \end{aligned}$$

- We define how different the filters are by taking the minimum of the KL divergences of all  $f_{k_i}$  and  $f_{k_j}$  pairs.
- The paper searches over binary patterns on  $13 \times 13$  grid that can be cut out of papers without having a floating piece. They randomly generate the patters and keep the one with the lowest KL score.

## 3 Deblurring

### 3.1 Deblurring with Maximum Likelihood Estimation

- For a fix filter  $f_k$ , we have that

$$\begin{aligned} \Pr(x|y) &\propto \Pr(x)\Pr(y|x) = \Pr(x)\Pr(n = y - C_{f_k}x) \\ &= \exp\left(-\frac{1}{2}x^*(\alpha C_{g_x}^* C_{g_x} + \alpha C_{g_y}^* C_{g_y})x\right) \exp\left(-\frac{1}{2}\frac{\|C_{f_k}x - y\|^2}{\eta^2}\right) \\ &= \exp\left(-\frac{1}{2}\left(\frac{\|C_{f_k}x - y\|^2}{\eta^2} + \alpha\|C_{g_x}x\|^2 + \alpha\|C_{g_y}x\|^2\right)\right). \end{aligned}$$

- We find  $x$  that maximize the above probability. That is, we would like to find  $\tilde{x}$  such that

$$\tilde{x} = \underset{x}{\operatorname{argmin}} \left( \frac{\|C_{f_k}x - y\|^2}{\eta^2} + \alpha\|C_{g_x}x\|^2 + \alpha\|C_{g_y}x\|^2 \right).$$

That is, we would like to find  $x$  that minimizes the reconstruction error  $\|C_{f_k}x - y\|^2$  with the constraint that the derivative of  $x$  should be small. (That is,  $x$  is smooth).

To minimize the above equation, we find the derivative with respect to  $x$ .

$$\begin{aligned} & \frac{\partial}{\partial x} \left( \frac{\|C_{f_k}x - y\|^2}{\eta^2} + \alpha\|C_{g_x}x\|^2 + \alpha\|C_{g_y}x\|^2 \right) \\ &= \frac{\partial}{\partial x} \left( \frac{\|C_{f_k}x - y\|^2}{\eta^2} \right) + \frac{\partial}{\partial x} \left( \alpha\|C_{g_x}x\|^2 \right) + \frac{\partial}{\partial x} \left( \alpha\|C_{g_y}x\|^2 \right) \\ &= \frac{1}{\eta^2} \frac{\partial}{\partial x} ((C_{f_k}x - y)^T (C_{f_k}x - y)) + \alpha \frac{\partial}{\partial x} (x^T C_{g_x}^T C_{g_x} x) + \alpha \frac{\partial}{\partial x} (x^T C_{g_y}^T C_{g_y} x) \\ &= \frac{1}{\eta^2} \frac{\partial}{\partial x} (x^T C_{f_k}^T C_{f_k} x - 2y^T C_{f_k} x + y^T y) + \alpha \frac{\partial}{\partial x} (x^T C_{g_x}^T C_{g_x} x) + \alpha \frac{\partial}{\partial x} (x^T C_{g_y}^T C_{g_y} x) \\ &= \frac{1}{\eta^2} \frac{\partial}{\partial x} (x^T C_{f_k}^T C_{f_k} x) - \frac{2}{\eta^2} \frac{\partial}{\partial x} (y^T C_{f_k} x) + \alpha \frac{\partial}{\partial x} (x^T C_{g_x}^T C_{g_x} x) + \alpha \frac{\partial}{\partial x} (x^T C_{g_y}^T C_{g_y} x). \end{aligned}$$

Using the properties that  $\partial(a^T x)/\partial x = a$  and  $\partial(x^T A x)/\partial x = (A + A^T)x$ , we have that

$$\begin{aligned} & \frac{\partial}{\partial x} \left( \frac{\|C_{f_k}x - y\|^2}{\eta^2} + \alpha\|C_{g_x}x\|^2 + \alpha\|C_{g_y}x\|^2 \right) \\ &= \frac{2}{\eta^2} C_{f_k}^T C_{f_k} x - \frac{2}{\eta^2} C_{f_k}^T y + 2\alpha C_{g_x}^T C_{g_x} x + 2\alpha C_{g_y}^T C_{g_y} x. \end{aligned}$$

Setting the above to 0, we have

$$\begin{aligned} & \frac{2}{\eta^2} C_{f_k}^T C_{f_k} x - \frac{2}{\eta^2} C_{f_k}^T y + 2\alpha C_{g_x}^T C_{g_x} x + 2\alpha C_{g_y}^T C_{g_y} x = 0 \\ & \frac{2}{\eta^2} C_{f_k}^T C_{f_k} x + 2\alpha C_{g_x}^T C_{g_x} x + 2\alpha C_{g_y}^T C_{g_y} x = \frac{2}{\eta^2} C_{f_k}^T y \\ & \left( \frac{1}{\eta^2} C_{f_k}^T C_{f_k} + \alpha C_{g_x}^T C_{g_x} + \alpha C_{g_y}^T C_{g_y} \right) x = \frac{1}{\eta^2} C_{f_k}^T y. \end{aligned}$$

So, we have to solve the linear system  $Ax = b$  with

$$\begin{aligned} A &= \frac{1}{\eta^2} C_{f_k}^T C_{f_k} + \alpha C_{g_x}^T C_{g_x} + \alpha C_{g_y}^T C_{g_y}, \text{ and} \\ b &= \frac{1}{\eta^2} C_{f_k}^T y. \end{aligned}$$

- Using this approaches introduces some artifacts along the edges, but this is not a concern in large image.
- However, the real drawback is that the estimated images tend to be overly smooth.

### 3.2 Deblurring by Non-Linear Optimization

- To produce sharper results, the paper changes the objective function to:

$$\|C_{f_k}x - y\| + \sum_{i,j} \left( \rho(x[i,j] - x[i+1,j]) + \rho(x[i,j] - x[i,j+1]) \right)$$

where the paper uses  $\rho(z) = |z|^{0.8}$ . This function is more heavy-tailed than  $|z|^2$  when thinking about it in terms of  $e^{-\rho(x)}$ .

Minimizing the above function still minimizes the reconstruction error. However, it encourages the derivatives to be more different, and thus results in sharper images.

- However, optimizing the above function is more complicated. The paper claims that this can be solved using the iterative reweighting least square (IRLS) process published in a subsequent paper [2]. (It's unclear to me how the above problem was turned into a problem where IRLS can be applied.)

### 3.3 Identifying the Correct Scale

- Let us assume that the whole picture's subject is a single depth. We can identify the scale corresponding to that depth by maximum likelihood estimation again:

$$\tilde{k} = \operatorname{argmax}_k P_k(y).$$

However, this is unreliable because of issues such as high-frequency noise in the image.

- Instead the paper tries to minimize the energy term  $E_k(y) = y^* \Sigma_k^{-1} y$ , which is the exponent of  $P_k(y)$ . Also, each scale's energy is scaled by the scale-specific constant  $\lambda_k$ , which was learned from training images to minimize error. That is,

$$\tilde{k} = \operatorname{argmax}_k \lambda_k E_k(y).$$

- Because  $y^* \Sigma_k^{-1} y$  takes a long time, the paper approximates it by

$$y^* \Sigma_k^{-1} y \approx \frac{1}{\eta^2} \|C_{f_k} \tilde{x} - y\|^2$$

where  $\tilde{x}$  is the deblurred image obtained from optimization.

## 4 Depth Determination

- Let there be  $K$  discrete depths  $k_1, k_2, \dots, k_K$ .
- The paper starts by computing the deblurred images  $x_{k_1}, \dots, x_{k_K}$  using each filter separately.
- Then, it computes the reconstruction error  $e_k = y - f_k * x_k$  for each depth  $k$ .
- For each pixel  $(i, j)$ , we approximate the error of using depth  $k$  in the neighborhood local to  $(i, j)$  by

$$\hat{E}_k[i, j] = \sum_{-s \leq \Delta i, \Delta j \leq s} (e_k[i + \Delta i, j + \Delta j])^2.$$

- The depth for pixel  $(i, j)$  is then the depth that minimizes the energy times the learnt constant in the last section:

$$d[i, j] = \operatorname{argmin}_k \lambda_k \hat{E}_k[i, j].$$

- The deblurred image is reconstructed as:

$$x[i, j] = x_{d[i, j]}[i, j].$$

- The depth estimate is often neither continuous nor smooth. For example, in the area where there's not much texture, any depth's filter will result in low reconstruction error.
- Instead, we seek a depth map that will be useful in applications such as refocusing after the fact and object extraction. This depth map must be (1) smooth in the area where the image is smooth, and (2) its discontinuities are aligned with image edges.
- The paper then casts this problem as an energy minimization in a Markov random field over the image. It seeks to find depth label  $\bar{d}$  over the image that minimize the energy function:

$$E(\bar{d}) = \sum_{i,j} E_1(\bar{d}[i,j]) - \lambda \sum_{i,j,i',j'} E_2(\bar{d}[i,j], \bar{d}[i',j'])$$

where

$$E_1(\bar{d}[i,j]) = \begin{cases} 0, & \text{if } \bar{d}[i,j] = d[i,j] \\ 1, & \text{if } \bar{d}[i,j] \neq d[i,j], \end{cases}$$

and

$$E_2(\bar{d}[i,j], \bar{d}[i',j']) = \begin{cases} 0, & \text{if } \bar{d}[i,j] = \bar{d}[i',j'] \\ \exp(-(y[i,j] - y[i',j'])^2 / \sigma^2), & \text{if } \bar{d}[i,j] \neq \bar{d}[i',j']. \end{cases}$$

This function is minimized using Boykov–Veksler–Zabih graph cuts algorithm [1].

- Sometimes, the depth labeling misses the exact layer boundaries due to insufficient image contrast. The paper incorporate user's brush stroke of required depth as additional input. The stroke is treated as hard constraint in the algorithm, and helps detect the right depth boundaries.

## References

- [1] Yuri Boykov, Olga Veksler, and Ramin Zabih. Fast approximate energy minimization via graph cuts. *IEEE Trans. Pattern Anal. Mach. Intell.*, 23(11):1222–1239, November 2001.
- [2] A. Levin and Y. Weiss. User assisted separation of reflections from a single image using a sparsity prior. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(9):1647–1654, sept. 2007.
- [3] Anat Levin, Rob Fergus, Frédo Durand, and William T. Freeman. Image and depth from a conventional camera with a coded aperture. In *ACM SIGGRAPH 2007 papers*, SIGGRAPH '07, New York, NY, USA, 2007. ACM.
- [4] Bruno A. Olshausen and David J. Field. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381(13):607–609, 1996.