Chair of Computer Science 6
RWTH Aachen University
Prof. Dr.-Ing. H. Ney

Aachen, 22. 01. 2006

D. Stein, R. Schlüter

Digital Processing of Speech and Image Signals

12. Exercise

Submission of the solutions: 29. 01. 2006 at the beginning of the lecture

---

Task 12.1 Template matching is a technique used in different image processing tasks. For example, variants of this technique could be used to detect objects for object-based video coding (cp. the MPEG standards).

Given a (square) template t of size $M \times M$, the goal is to find the position in the given (square) image x of size $N \times N$, $N > M$ that resembles most the given template. Using the Euclidean distance as a measure of (dis)similarity, we are looking for:

$$(\hat{n}_1, \hat{n}_2) = \text{arg} \min_{n_1,n_2} \sum_{m_1=1}^{M} \sum_{m_2=1}^{M} \left[\, x(n_1+m_1, n_2+m_2) - t(m_1, m_2) \,\right]^2 \tag{1}$$

a) Rewrite the squared difference using the binomial formula $(a-b)^2 = a^2 - 2ab + b^2$ and convert the summation into three separate summations. Which of the resulting terms can be neglected for the minimization as it does not depend on (n_1, n_2)? (1 P.)

b) Continuing with your result of part (a), show that the sum involving x and t (as a function of (n_1, n_2)) is a convolution of x and an appropriate modification t' of t. Give t' as a function of t. (1 P.)

c) Denote the sum of part (b) with $C(n_1, n_2)$. How can you efficiently determine arg $\max\limits_{n_1,n_2} C(n_1, n_2)$ using the FFT? (1 P.)

d) Write a function to determine arg $\max\limits_{n_1,n_2} C(n_1, n_2)$ for the image `office` and the template `template` (given on the homepage of the lecture). Visualize C as an image and interpret the result. (3 P.)

e) The remaining third summation can be efficiently calculated using running sums. In one pass over the image you can calculate

$$S(n_1, n_2) = \sum_{n=1}^{n_1} \sum_{n'=1}^{n_2} x^2(n, n')$$

for all $n_1, n_2 \in \{1, \ldots, N\}$. How? How can you efficiently calculate sums of squares of the type

$$S(n_1, m_1, n_2, m_2) = \sum_{n=n_1}^{m_1} \sum_{n'=n_1}^{m_2} x^2(n, n')$$

using the table S? (1 P.)

f) Use your results from parts (a) to (e) to give an algorithms to determine $(\hat{n}_1, \hat{n}_2)$ in complexity $O(N^2 \log N)$. Give the result for the image `office` and the template `template` (given on the homepage of the lecture). What would the complexity of a naive implementation of equation (1) be? In which case would the naive implementation be preferable? (3 P.)

Dennis Gabor (1900-1979) was a Hungarian physicist and inventor who is most notable for inventing holography. At the start of his career, he analyzed the properties of high voltage electric transmission lines by using cathode-beam oscillographs, which led to his interest in electron optics. Having fled from Nazi Germany in 1933, Gabor was invited to Britain to work at the development department of the British Thomson-Houston company in Rugby, Warwickshire. While working there, he invented holography in 1947, an achievement for which he later received the Nobel Prize in Physics in 1971.
In *The Mature Society : a view of the future* in 1972, Gabor wrote: *The best way to predict the future is to invent it.*
Source: `http://www.wikipedia.com`