Photogrammetrische Woche 1989
Institut für Photogrammetrie
Schriftenreihe, Heft 13, 1989

DIGITAL PROCESSING OF HIGH DEFINITION IMAGES

By Dr.-Ing. Rudolf E. Grosskopf, Oberkochen

Data Acquisition

To process images, acquiring the image data is an important prerequisite. Therefore the first topics to be discussed are how to acquire and store high definition images with digital means.

When Armstrong transmitted the first pictures from the moon with a hand-held camera, the electronic sensor technology was only slightly better than home TV. This situation has in the meantime changed dramatically. In studio applications, CCD sensors resolve more than 1,000 lines. If today's still very high prices can be paid, even images with more than 2,000 x 2,000 = 4 x 106 pixels can be acquired with one CCD sensor frame.

However, in photogrammetric applications, the traditional technology is much more powerful. Highly resolving objectives and fine-grain photo emulsions yield images with up to 1.2 GBytes per frame at very moderate cost. State-of-the-art cameras are equipped with precision shutter systems and forward motion control. So, for the time being, it may be said that electronic imaging is still some orders of magnitude away from the state of the art achieved in photogrammetry with conventional means. One reason may be the long development time. Photogrammetry began at the beginning of our century already to perform powerful image processing for mapping.

Does this mean that the technology for electronic image processing is still very far behind the needs of photogrammetry?

In order to get the right impression of the currently available technology for acquiring and storing high definition images, another point can be made. The amount of data which has to be acquired in a given time has to considered. Figure 1 gives an overview. The resolution of a standard TV image is 0.8 MBytes per frame in standard TV systems. Slides can contain 20 MBytes and cost less than $1.00. The summit of data content is reached in 23 cm x 23 cm frames exposed with photogrammetric cameras. Up to 1.2 GBytes can be stored in one single photo.

In its lower part, Figure 1 gives some relevant information answering the question. Two numbers are given about the data throughput within 10 seconds in photogrammetric cameras and TV cameras. During automatic operation, the photogrammetric camera exposes approximately one image every 10 seconds. In the same time a TV camera takes 500 to 600 frames of a vivid scene. Between 200 and 400 MBytes of information capacity are needed to store these images digitally. Compared with the 1.2 GBytes in a photogrammetric slide it is the same order of magnitude.

So it may be concluded that the electronic technology is already capable of handling the signal throughput that is needed for high definition images. But the existing technology has not yet been adapted to the specific needs of photogrammetry. An appropriate effort for photogrammetric applications will probably result in a deeper penetration of this field by electronics.

A very recent innovation confirms this conclusion. It is the camera with programmable resolution by Reimar Lenz. He had a very effective idea and implemented it. Figure 2 shows the principle of his camera. A CCD sensor with an array of diodes is mounted on two piezo actuators, one for the x direction and the other for the y direction. The light sensitive diodes on the chip only partially cover the chip area. This drawback of some chip designs has to be accepted and is turned into a benefit by Lenz's camera design. The whole sensor chip is moved stepwise in the indicated way. For each position one frame of pixels is digitized. Thus electronic imaging with a CCD sensor yields 20 MBytes at reasonable cost. Compared with the .8 MBytes of a good TV camera, this is an improvement by the factor 25.

Figure 3 visualizes this progress more vividly. Figure 3a shows the full frame of a scene. Figure 3b is an enlarged segment of that scene with the resolution of a good home video camera. Figure 3c shows the resolution of a TV studio CCD colour camera. Figure 3d demonstrates the much better resolution of the ProGress camera by Reimar Lenz.

Even more suited for photogrammetric applications is the time delay integration (TDI) principle. Chips exploiting this method on silicon are available from different vendors. TDI helps to overcome the signal to noise ratio problem that exists in airborne line scanning cameras. Such cameras only utilize the light of one pixel line at a given time. The rest of the field of view of the objective is not used. Compared with photographic film, which takes advantage of the whole field of view, this means wasting more than 99.9 % of the light energy.

Time delay integration with currently available TDI sensors exploits the light on 128 lines simultaneously. The chip shown in Fig. 4 is supplied with two types of clock. The shift clock is used to transfer the analog information of one TV line. This happens in the traditional way of line scanners, pixel after pixel. The TDI clock is TDI-specific. It is used to transfer the accumulated charges perpendicularly to the line scanning direction, i. e. from one line to the next until it reaches the readout line.

The operation resembles the movement of the photographic film in a forward motion control camera. The movement of the film plays the same role as the charge transfer between the pixel lines in TDI sensors.

During operation, the time delay clock of TDI sensors has to be matched to the aircraft speed. The means for that can be similar to those used for forward motion compensation.

The uniform image motion can be the movement of a sheet of paper in a document scanner or the movement of an aircraft in photogrammetric applications. As indicated, the sensor has 128 TDI lines. The TDI clock has to be matched to the movement of the object so that the accumulated charges are transfered with the speed of the image on the sensor. The accumulated charges are transferred from the last line via shift registers to two outputs: one at the left side, the other at the right side of the chip. Each output transfers 515 of the 1030 pixels.

At present, sensors with as many pixels per line as photogrammers need are not yet available. However, it can be assumed that by stacking many such sensors or by developing longer versions it should be possible to scan also with the resolution of traditional photogrammetric cameras.

Data Storage

If, in future cameras, digital image data came as a stream of e. g. 120 MBytes per sec., system designers may ask how to store them. It may be impossible to store that much digital data in a system with a reasonable size, weight and price for airborne systems.

Probably the biggest existing system for storing digital data is the WURLITZER by Kodak. It is used as a backup data storage in mainframe computers and contains 150 optical discs in a juke-box like system with 6.8 GBytes per disc. It could store 900 images with 1.2 GBytes each. However, this system is so large and the price so high that it is difficult to believe in this principle for photogrammetric applications. Even if a radio-frequency data link were used between the airborne camera and a stationary WURLITZER to eliminate the need of installation aboard the aircraft, the WURLITZER would be too cumbersome.

Another idea offers more hope for a solution of the digital storage problem. Just recently the technological leaders in digital TV imaging developed the technology to transfer TV images with only 64 KBits per sec. while 216 MBits per sec. are needed with conventional means. This data compression by coding is done with the DCT (discrete cosine transform) method /Gerhard 88/. A compression factor of 3000 for teleconferencing applications has been achieved after many years of research and development. Scientists eventually came up with the right algorithm and the right silicon chips to build practical systems which are being tested now.

It is unlikely that DCT will also meet the high geometric precision requirements of photogrammetry. However, the quality of DCT-compressed, transferred and re-expanded images shows that the information which the human eye wants to see on the monitor is deteriorated only little by compression. So it might be possible to find a compression method that is suited to the compression of geometrically precise data.

Processing the Geometric Information in Images

Convolution

Convolution is a method which proved useful in many applications in signal processing and also in metrology. To facilitate understanding, the principle is described first.

Figure 5a demonstrates the convolution principle in an application for smoothing a noisy signal. Signal f is convoluted with rectangle g. The result is the smoothed signal h. This process is equivalent to computing the mean value of that part of f that is windowed by the rectangle, and of shifting the window continously along the abscissa of f.

Convolution can also be used for edge enhancement. A modified function g as shown in Figure 5b can be used for this. It begins with a small negative amplitude. A higher positive peak follows and a small negative deflection terminates the function g. When a signal f with a smooth slope is convoluted with this shape of g, a steeper slope with overshoots results.

To understand the process of convolution of one function f with another function g, as mentioned previously, one may think of moving the function g in the x direction over the function f and of multipling all values of both functions for e a c h position. Furthermore, the products for each position of g have to be integrated. Of course, the number of values that have to be computed will be very high if the functions contain much detail (i. e. high frequencies).

To illustrate more vividly the amount of computation needed for convolution in image processing, a practical example is helpful: a standard state-of-the-art image has 512 x 512 pixels with 8 bits each. That is .25 MBytes. If the largest pattern to be recognized covers 8 x 8 = 64 pixels in the image, 64 multiplications and additions (multiply/adds) have to be computed for every possible position of the window in the image. Therefore 84,000,000 multiplications and additions are required to convolute online the images of a 50 frames per sec. camera using the 8 x 8 pixel pattern.

It is one of the amazing events of the last 20 years that engineers were able to work out a technology which yields circuits meeting these needs. This result was accomplished by the interaction of optics, fine mechanics and physics. As a result, many algorithms and also convolution have become practicable. Convolution is now one of the most powerful tools in signal and image processing. Chips are available which cost only a few hundred dollars and handle more than one billion instructions per second.

Compression

As explained above, it may make sense to restrict the storage and also the computation needs to as little data as possible. Data compression remains useful also in the world of powerful megaflop and megabyte chips and high-performance computers because the amount of data which current and future sensors can put out exceeds the storage and computing capacity of even the most powerful hardware.

There are many methods to decompose images into subsets with reduced amounts of data. Fourier transform, cosine transform, Hadamard transform, Haar treansform, Karhunen-Loeve transform and Singular Value Decomposition (SVD) are only some of them /Pratt 78/. For every transform, examples of images can be found for which they are ideally suited, and others for which they are totally inappropriate. To describe them all is not possible here. Instead, only SVD is explained. We used it successfully for data compression in order to reduce the amount of computations needed for position measurement.

SVD

Singular value decompostion is one of the achievements of matrix calculus. It can be used to reduce the amount of data and of computations needed in practical applications. Contrary to the above-mentioned DCT it promises not to impair the geometric precision. This assertion will be proven by an example.

Figure 6 shows in principle what singular value decomposition does. The original of a car is represented by 65,536 1-byte pixels in the upper image. Singular value decomposition computes vectors defining grey level distributions in the x and y directions. The x and y vectors for n = 1 already show that a concentration of darkness is located near the rear wheel of the car in the lower left part of the image.

Using six vectors (n = 3) yields a pattern that shows a shape similar to a car. With n = 20 (40 vectors) the car is almost perfectly restored.

Although the vectors of singular value decomposition generate only linear patterns in the x direction and other linear patterns in the y direction, circular structures like the wheels are also reproduced.

The mathematics behind this are omitted here. If you are interested, you might read any book about matrix computations, e.g. /Golub 83/.

As shown by the above example, two vectors of SVD already represent specific information on the object. The step from the 64 KBytes of the full image to two vectors with only 256 bytes each yields a compression factor of 128.

Now the question has to be answered if the two vectors already represent enough information to determine the position of a specific pattern.

This question is answered by means of the practical problem of measuring the position of holes in mechanical parts. Figure 7a shows the original image at the left. The two holes in the two lugs are to be identified automatically and their distance measured. For that purpose, a Sobel enhancement is first used to accentuate the edges of the specimen. Also, a region of interest is defined as a rectangular area. The Sobel operation outperforms a specific version of 2D differentation in image processing /Haberäcker 87/. Then a template as shown in the left part of Figure 7b is taken. This template is the Sobel enhanced upper hole. Now SVD comes into play. The middle picture in 7b shows the reduced version of the template (n = 1). The image at the right shows the same template with n = 2.

As indicated above, every convolution yields a curve for 1D processes and a two-dimensional set of data for images. The 2D set of data can also be displayed as an image. This is done in Figure 7c. In the left part it shows the convolution result of the specimen with the three versions of templates.

The left image shows convolution with the full template. Looking at the numerical data not shown here it can be seen that there are three maxima. One for the upper hole, one for the piece of dirt between the lugs, and the third for the lower hole. The maxima for the upper and the lower holes are the highest maxima.

When using only the reduced template with n = 1, very similar data results. The upper and lower holes again cause the highest maxima.

In practical applications, the positions of the maxima are used to determine the position of the holes. For n = 1 the position error is 3 pixels, which is too high. With n = 2 the error shrinks to less than .5 pixels, which is good enough for many practical applications.

The full template is made up of 2,500 bytes. The reduced template n = 2 contains only 200 bytes. This is a reduction factor of 12. So it can be said that the example shows the practical applicability of singular value decomposition to reduce the amount of required data and processing. Despite a reduction factor of 12 the geometric precision contained in the original data is retained successfully.

As the example shows, singular value decomposition is a suitable mathematical method for reducing the amount of data needed for measuring the position of holes. It has been proven that it works with non-ideal data from a scene with limited contrast and imperfect illumination. The achieved data compression exceeds the factor 10. This may prompt a research project to define a specific method for compressing digital data in airborne digitizing cameras. Of course, such a project should also consider more traditional transform methods such as DCT, Hough transform, Fourier transform, run length coding and others.

Summary

The data content of high definition images is an important aspect in digital processing. Albeit the electronics for TV applications proves that the amount of data can be handled in principle, the appropriate mathematics and basic algorithms have not yet been investigated. It has been shown, however, that powerful data compression methods exist. They are useful for applications such as digital TV systems (teleconferencing) and metrology. The algorithms for image enhancement and pattern recognition also exist (e. g. averaging, convolution and singular value decomposition). Combining such compression and processing methods in appropriately structured software will probably lead to a deeper penetration of high definition image processing by digital electronics.

Acknowledgements

I want to thank Dr. Klaus Knupfer for helpful discussions, Dipl.-Ing. Klaus Suedland for contributing images computed with the singular value decomposition method, and Dr. Manfred Schmutz for the SVD software.

REFERENCES

/Gerhard 88/ Alexander Gerhard, Bewegungsanalyse bei der Codierung von Bildsequenzen, Dissertation München 1988

/Castleman 79/ Kenneth R. Castleman, Digital Image Processing, Prentice Hall N. J.

/Golub 87/ G. H. Golub, C. S. Van Loan, "Matrix Computations", Johns Hopkins University Press, Baltimore MA, 1983.

/Haberäcker 87/ Haberäcker Peter, Digitale Bildverarbeitung, Hanser, München

/Pratt 78/ William K. Pratt, Digital Image Processing, Wiley and Sons

ABSTRACT

An attempt is made to assess existing technologies for using digital image processing ing in the entire spectrum of information processing in photogrammetry. First the amount of data supplied by photogrammetric cameras is compared with standard image processing applications. The amount of data in photogrammetric images to be stored digitally is studied and it is shown that the present state of the art of digital storage means does not meet photogrammers' needs. Examples are given of successful data compression in other fields of application of digital technology. They show that there may be a way to compress digital data from photogrammetric images and thus to open up a way for online storage of the images. Singular value decomposition is shown to be a suitable mathematical method for compressing geometric data without impairing the geometric precision. Convolution is explained and its computation requirements are discussed. The use of SVD is illustrated by means of an application taken from metrology.

TECHNIKEN DER DIGITALEN VERARBEITUNG HOCHAUFGELÖSTER BILDER

ZUSAMMENFASSUNG

Es wurde versucht, die technologischen Voraussetzungen zu bewerten, die zu erfüllen sind, um digitale Techniken im gesamten Spektrum der Aufgaben in der Photogrammetrie einzusetzen. Zunächst wird das Datenvolumen photogrammetrischer Aufnahmen im Vergleich zu Standardsystemen der Bildgebung betrachtet und es wird erläutert, daß der gegenwärtige Stand der Technik bei der digitalen Speicherung die Bedürfnisse der Photogrammetrie nicht erfüllt. Beispiele erfolgreicher Datenkompression aus anderen anderen Anwendungsgebieten werden erwähnt. Sie zeigen, daß es einen Weg geben kann, um den Datenstrom aus digitalisierten photogrammetrischen Bildern so zu komprimieren, daß die Bilder online abgespeichert werden können. Die Singular Value Decomposition (SVD) wird an einem Beispiel als Verfahren aufgezeigt, das Bilddaten in einer Weise komprimiert, die deren geometrische Präzision nicht nennenswert beeinträchtigt. Faltung wird erläutert und die Hardwarevoraussetzungen für videoschnelles Falten werden quantifiziert. Ein Beispiel der Anwendung der SVD auf ein mechanisches Teil wird beschrieben.

Dr.-Ing. Rudolf E. Grosskopf
Carl Zeiss
Zentralbereich Forschung
Forschung Elektronik und Informatik
Elektronik-Koordination
Carl-Zeiss-Straße

7082 Oberkochen

Photogrammetrische Woche 1989
Institut für Photogrammetrie
Schriftenreihe, Heft 13, 1989

slide
20 Mbyte

TV - frame
.8 Mbyte

photogrammetic image
1.2 Gbyte

Resolution of Images

photogrammetic camera
1.2 Gbyte

TV - camera
.4 Gbyte

Throughput in 10s

Fig. 1 Amount of Data

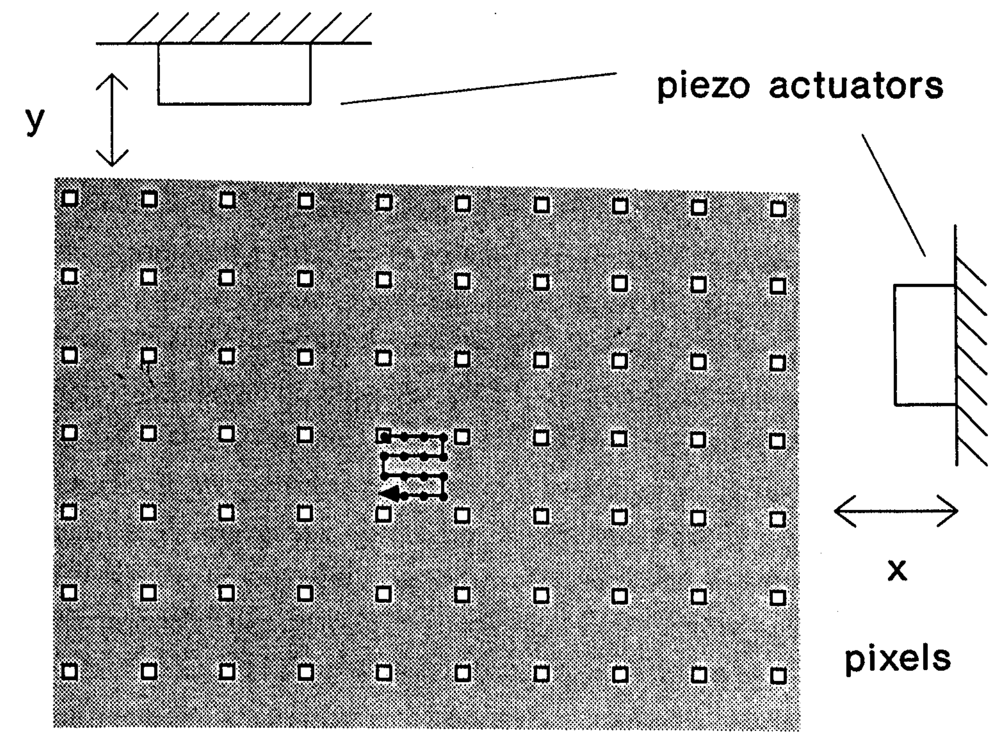
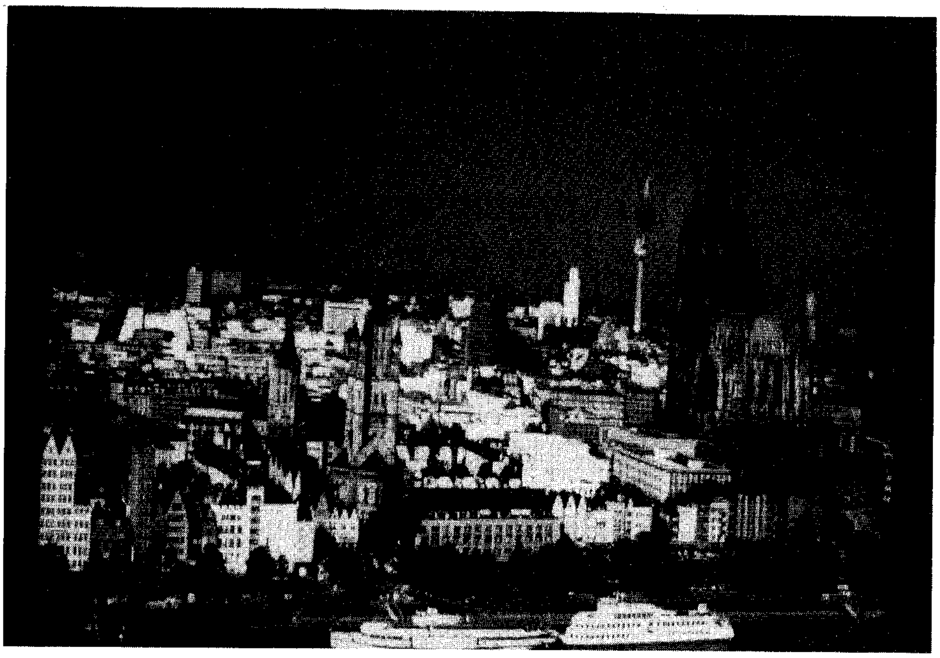


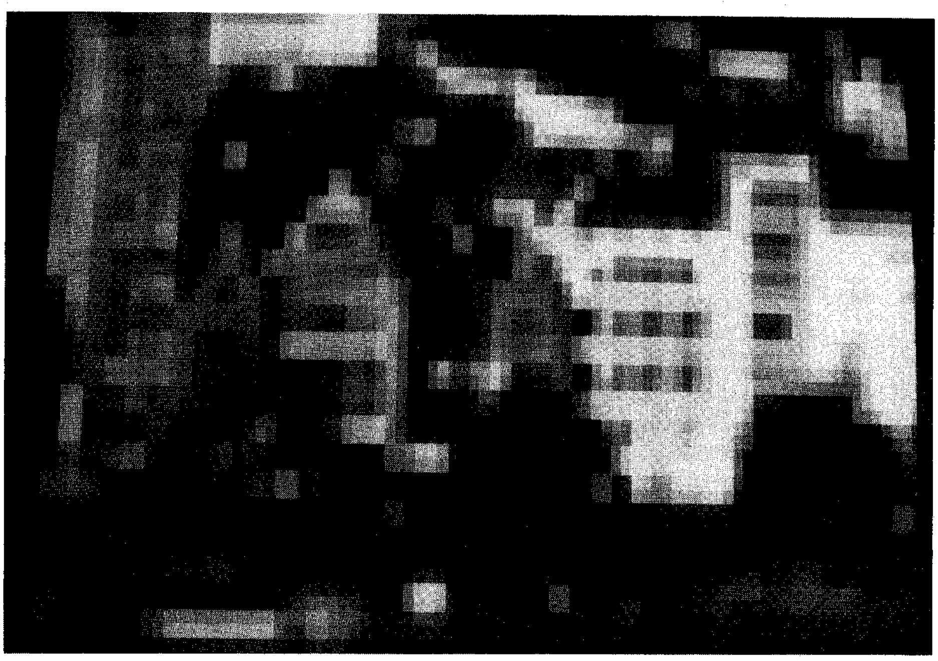
Fig. 2 Camera for Still Images

(adapted from REIMAR LENZ)

Photogrammetrische Woche 1989
Institut für Photogrammetrie
Schriftenreihe, Heft 13, 1989



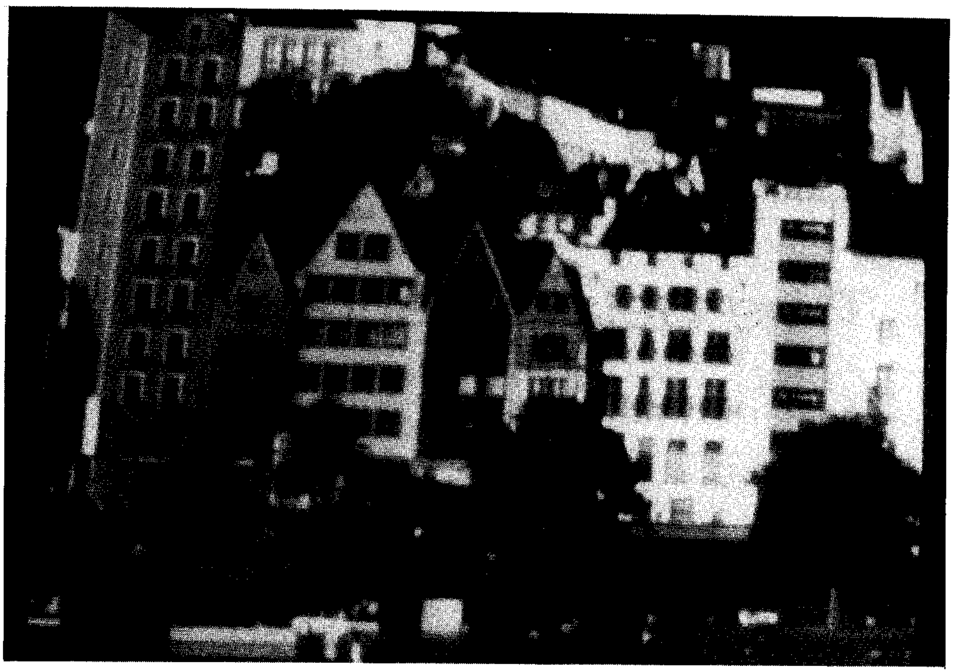
a) full frame



b) enlarged segment from a) with the resolution of a good consumer video camera

Fig. 3 a, b Resolution Comparison
(adapted from Reimar Lenz)

c) segment with the resolution of a TV-Studio CCD-Colorcamera three chips with 750 x 580 pixels each



d) segment with the resolution of the Camera by Reimar Lenz

Fig. 3 c, d Resolution Comparison
(adapted from Reimar Lenz)

Photogrammetrische Woche 1989
Institut für Photogrammetrie
Schriftenreihe, Heft 13, 1989

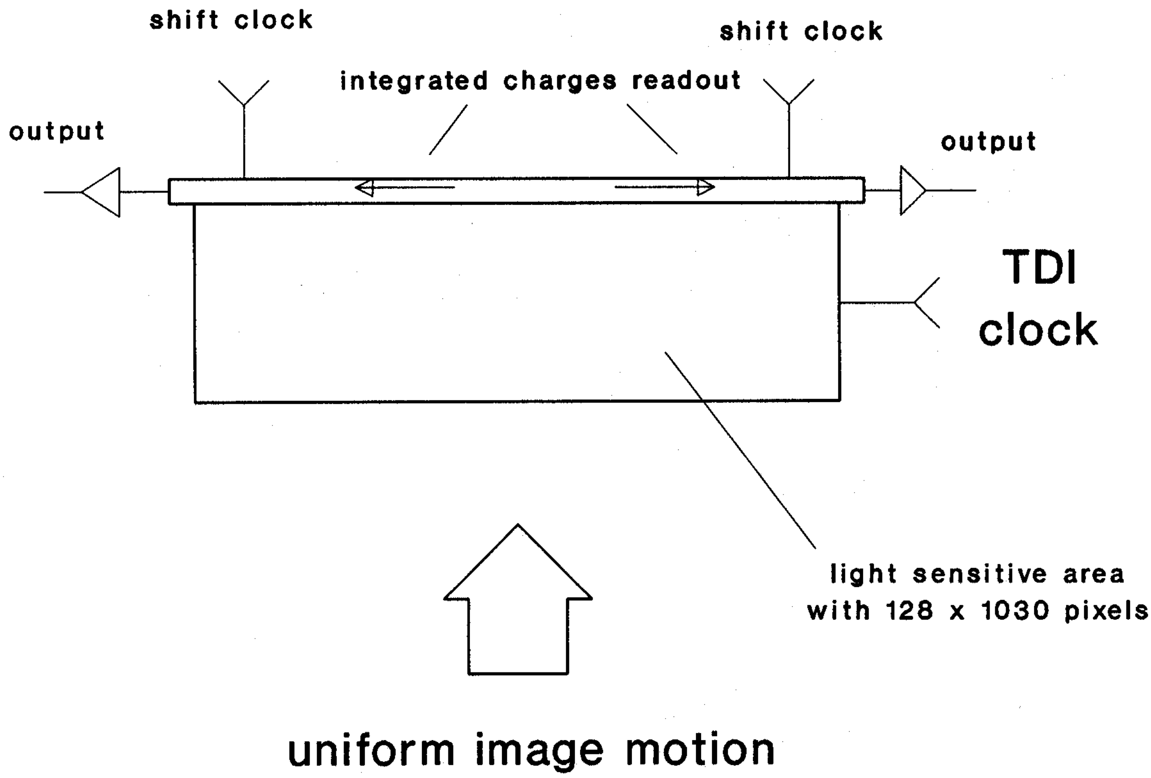


Fig. 4 Time Delay Integration

Photogrammetrische Woche 1989
Institut für Photogrammetrie
Schriftenreihe, Heft 13, 1989

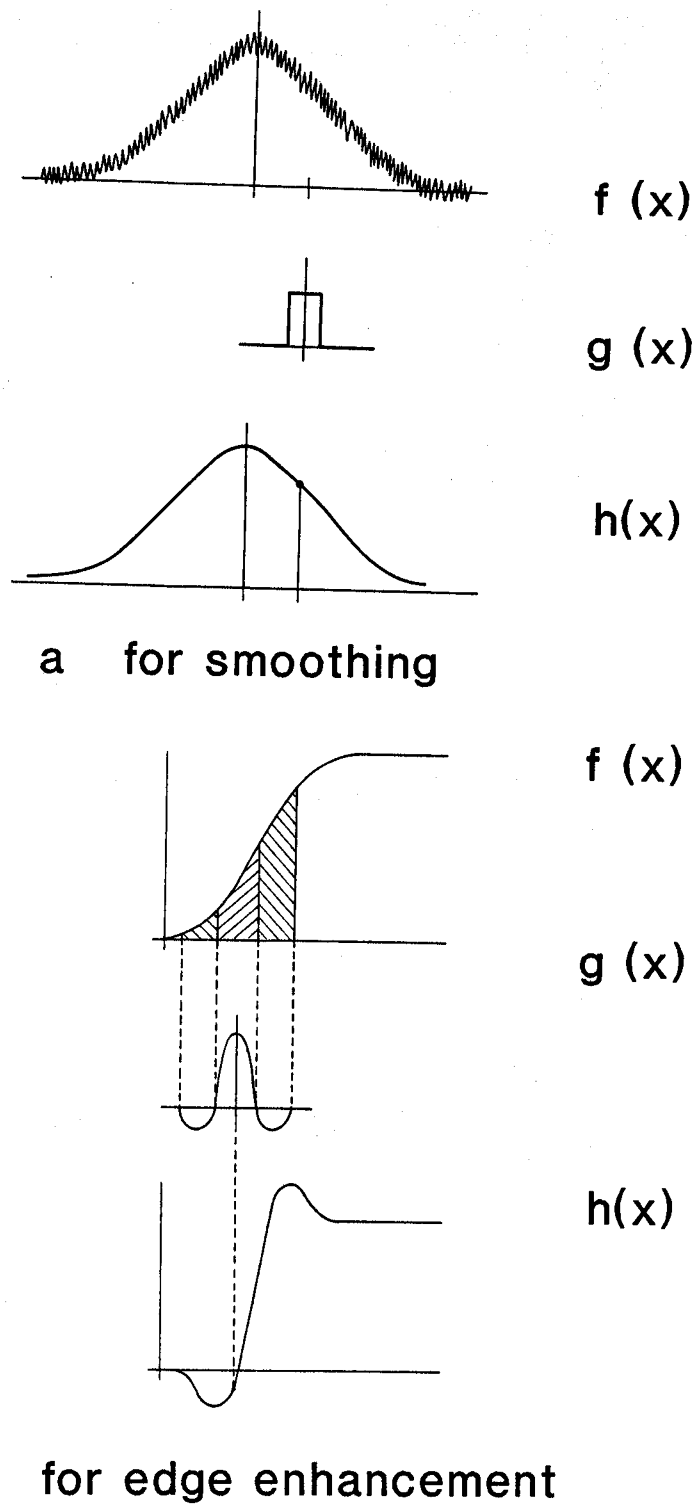
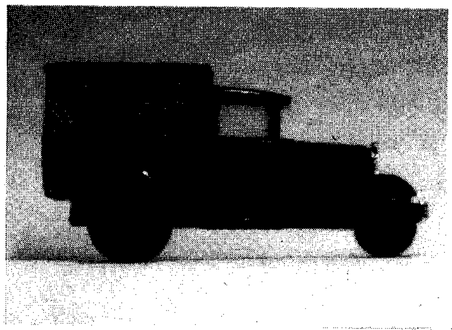


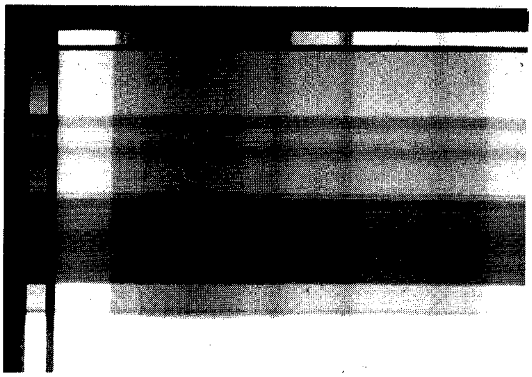
Fig. 5 Convolution

(adopted from Kenneth R. Castleman, Digital Image Processing, Prentice Hall N. J.)

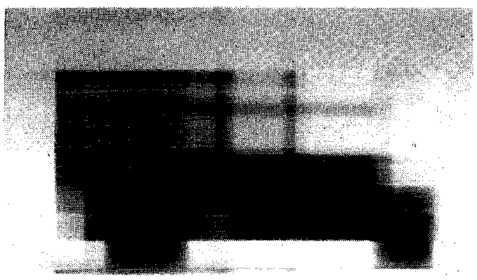
Photogrammetrische Woche 1989
Institut für Photogrammetrie
Schriftenreihe, Heft 13, 1989



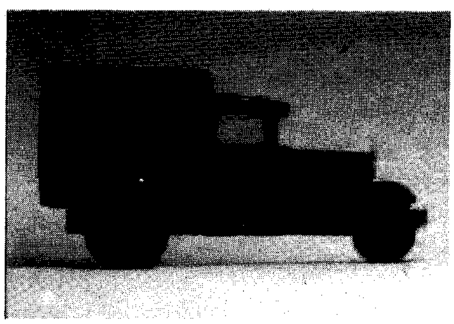
Original
(65 536 bytes)



n = 1
(2 vectors, 512 bytes)

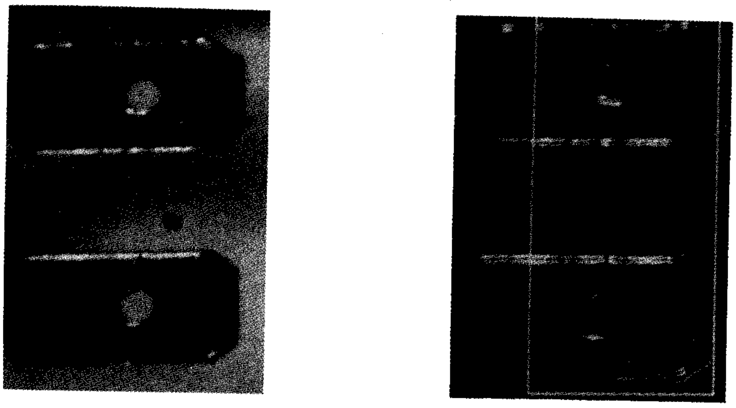


n = 3
(1536 bytes)



n = 20
(10240 bytes)

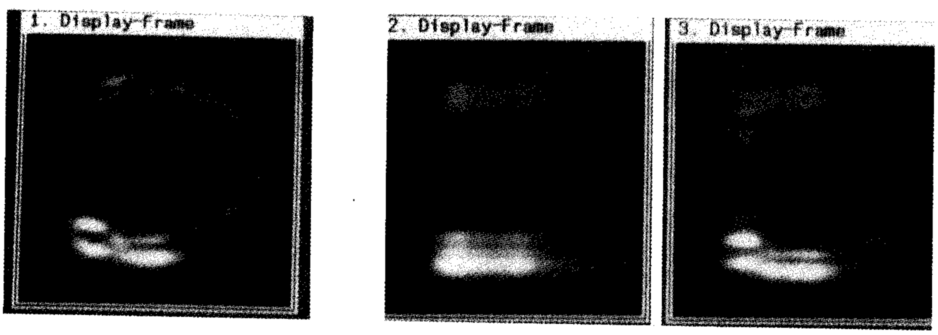
Fig. 6 Singular Value Decomposition
(S V D)



original

Zobel enhanced

a

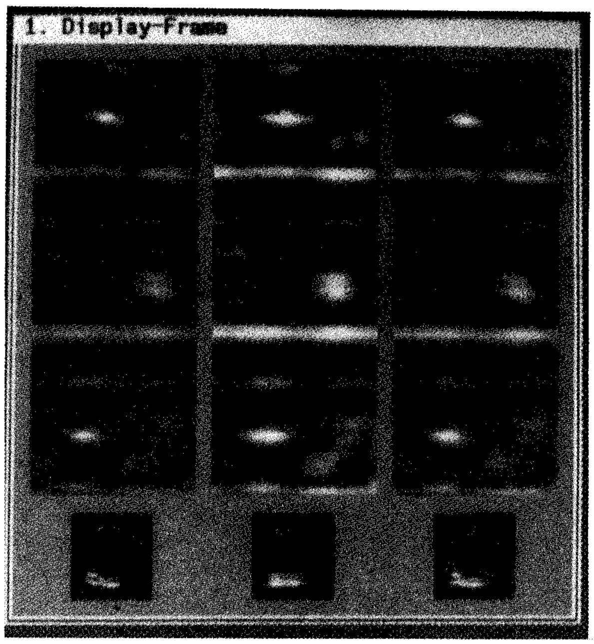


full template
(upper hole from a)

SVD
n = 1

SVD
n = 2

b



convolution results

left: full template
middle: SVD-template n = 1
right: SVD-template n = 2

c

templates used
repeated from b

Fig. 7 Position Measurement of Holes