CANADIAN THESES ON MICROFICHE

THÈSES CANADIENNES SUR MICROFICHE

National Library of Canada
Collections Development Branch

Canadian Theses on
Microfiche Service

Ottawa, Canada
K1A 0N4

Bibliothèque nationale du Canada
Direction du développement des collections

Service des thèses canadiennes
sur microfiche

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30. Please read the authorization forms which accompany this thesis.

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30. Veuillez prendre connaissance des formules d'autorisation qui accompagnent cette thèse.

THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED

LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE

Canada

NL 339 (r. 85/09)

Real Time Motion Detection in Digital Subtractive Angiography

by

Harry, Cheng Ho, Oung

A thesis
presented to the University of Ottawa
in partial fulfillment of the
requirements for the degree of
Master of Applied Science
in
Department of Electrical Engineering

OTTAWA, Ontario, 1985

UNIVERSITÉ D'OTTAWA
UNIVERSITY OF OTTAWA

The University of Ottawa requires the signatures of all persons using or photocopying this thesis. Please sign below, and give address and date.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my thesis advisor, Dr. Andrew Smith, for his guidance throughout this research. His suggestions and support have been extremely helpful and genuine. Under his supervision, the organization of my thought has greatly improved. The considerable help from Dr. Morris Goldberg is also greatly appreciated, especially in the part of literature review.

In addition, I wish to thank my colleague Sun Hui Fang for his encouragement and his various comments. Finally, the financial support from the NSERC grant is also acknowledged.

PREFACE

In the recent years, the applications of digital image processing in medicine have increased at a tremendous rate. Some well known diagnostic imaging modalities based upon digital computer technology for data acquisition and image production are computer tomography, nuclear magnetic resonance and ultrasound. Recently, the technique of Digital Subtractive Angiography (DSA) is finding widespread application in "intravenous angiography".

Prior to its development, all visualization of arteries in the body was performed with a high risk, time consuming procedure involving "intra-arterial injection" of contrast medium. DSA enables the performance of a real time, less dangerous procedure utilizing the image enhancement technique of subtraction where unwanted background information from an x-ray image is removed with the subtraction of two images taken before (mask) and after (live) a contrast agent has been introduced. For meaningful subtraction, the technique requires that patient motion does not occur and that the iodinated vessels themselves do not move during or between the acquisition of the images. However, neither of these is universally true in clinical DSA practice. Presently, the solution to this motion problem is to repeat the

image aquisition procedure. In so doing, both patient and physician will be exposed to radiation which would not have been necessary. It is, therefore, the purpose of this dissertation to develop a real time approach to counteract the effect of motion and to study its effectiveness.

In this thesis, a new algorithm called "the overall contrast of the difference image (OCD)" is described and explored analytically and experimentally. In Chapter 1, there is a general introduction to DSA covering the basic principle of x-ray subtraction and the engineering aspect of the DSA system. The presentation provides the reader with an understanding of the general physical and technical requirements of DSA and the problem of patient motion. Chapter 2 consists of a brief analysis of motions that could occur in a clinical examination and an overview of various modifications to DSA.

Chapter 3 presents a real time remasking scheme which requires only an additional motion detector unit to the DSA system. With this approach, motion artifacts are minimized by updating the mask image whenever a significant motion is detected such that the adverse effects of patient motion can be reduced or eliminated. Also included in this chapter is an overview and analysis of various image processing approaches to motion estimation. Based on a mathematical analysis, a new approach to motion estimation using image

statistics which could be obtained in real time is developed in Chapter 4. In Chapter 5, the approach presented is performed upon clinically acquired angiogram images and finally, in Chapter 6, comments related to the test and conclusions are given.

CONTENTS

Chapter I

DIGITAL SUBTRACTIVE ANGIOGRAPHY (DSA)

The field of medical imaging has been growing rapidly since the discovery of x-rays in the late 19-century and has had a great impact upon the practice of medicine. With the advances in digital computer technology, new diagnostic imaging modalities are being developed. Some well-known examples are computer tomography, ultrasound and nuclear magnetic resonance.

Recently, digital subtractive angiography is emerging as a useful tool for blood vessel imaging. The imaging of blood vessels, especially the arteries that supply blood to vital organs such as the brain and the heart, provides important information for patient diagnosis. Occlusions of the arterial system cause heart attacks, strokes and other serious medical problems. Together, these represent the major cause of mortality in our adult population. The discovery of vessel blockage or narrowing, associated with the exact location of the disorder, therefore, constitute vital parts of medical diagnosis.

1.1 BLOOD VESSEL IMAGING (ANGIOGRAPHY)

In conventional Radiography, X-Rays are directed through the examined body to produce images on photograghic plates or fluoresent screens. Due to the difference in X-ray attenuation of different organs and tissues, X-ray images outline the anatomical details in discernible contrast. Unfortunately, because the differential X-ray attenuation between blood and the surrounding soft tissues is so small, blood vessels are generally not visible.

1.1.1 Contrast Medium Injections

In order to visualize blood vessels, it is necessary to introduce a radio-opaque contrast agent or dye (a water soluble compound containing iodine). In conventional angiography (X-ray examination of blood vessels), a small polyethylene tube (catheter) is inserted directly into the vessel of interest. A high concentration of contrast agent is then injected. The resulting images display a vessel contrast large enough to be seen without further processing. However, in the case of arteriography (visualization of arteries), the requirement that the catheter be inserted into the artery increases the risk of the procedure, and hospitalization of patient is necessary [6,46].

If the injection is made into a vein (intravenous injection), and the images taken when the bolus of contrast agent

has traversed the circulation system to the artery, the artery would be poorly visualized. This is because by the time the contrast agent reaches the artery of interest, it will be so dilute that the contrast is too small for visualization. However, such a procedure would be of great benefit because of the relative safety of intravenous injection. It remains therefore to find ways of enhancing the low vessel contrast.

Subtraction is an imaging technique that provides low contrast signal enhancement, by removing the high contrast background structure (eg. soft-tissue, bone) from the images. In the process, X-ray images taken before and after the contrast injection are subtracted so that any change in the images (such as the opacification of vessels from contrast agent) is enhanced. Clinically, valuable images of arteries can therefore be obtained with a diluted iodine concentration (2%-3% contrast in the blood as opposed to that of 40%-50% in arterial catheter angiograms [12]). In this connection, intravenous angiography is made possible avoiding certain risks, much cost and hospitalization of patient.

1.2 THE PRINCIPLES OF IMAGE SUBTRACTION

1.2.1 The Classical Method

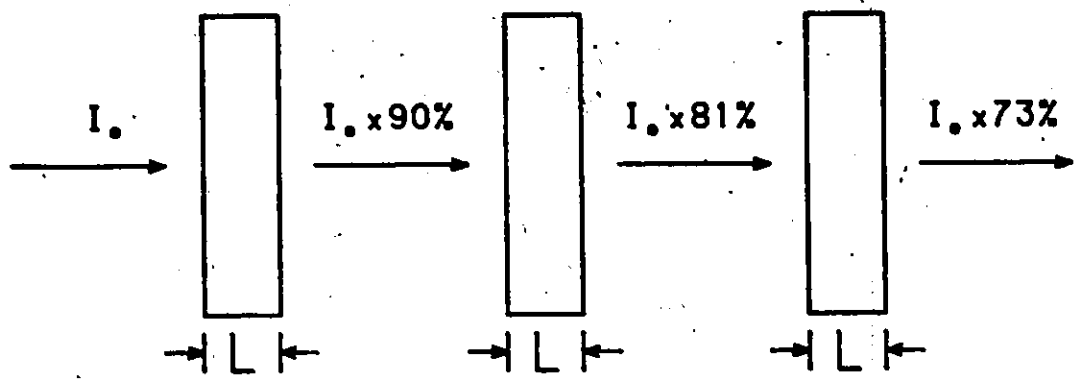
In a radiographic system, the X-ray energy available to the recording medium is attenuated exponentially as it passes through the patient body or specimen of interest (Figure 1). Assuming that X-rays propagate in the z-direction, and neglecting scattering and diffraction, the transmitted intensity available to the recording medium [10] is given by

$$I(x,y)=I_o \exp[-\int_o^L \alpha(x,y,z)dz]$$
$$=I_o \exp[-\beta(x,y)] \qquad (1)$$

where I_o is the spatially uniform incident radiation
L is the thickness of the specimen
$\beta(x,y)$ is defined as the attenuation projection
and $\alpha(x,y,z)$ is the linear attenuation coefficient at point (x,y,z) of the object

The process for low contrast enhancement involves the subtraction of two radiographic images taken before (mask) and after (live) a suitable contrast agent has been introduced. If the mask image is denoted by $I_m(x,y)$, the attenuation coefficient of contrast agent by μ, iodine concentration by $c(x,y)$ and the spatial distribution of vessel thickness by $t(x,y)$, then the intensity of the image containing opacified contrast agent is given by

$$I_l(x,y)=I_m(x,y)\exp[-\mu c t(x,y)] \qquad (2)$$



$$\frac{dI}{dz} = kI \quad ===> \quad I = I_o \exp(kz) = I_o \exp(-\alpha z)$$

Figure 1: Exponential attenuation of Radiation through Matter. The absorption of x-rays in a thin layer of substance is proportional to the thickness of the layer and the amount incident on that layer.

where c is assumed to be space invariant.

The difference projection, which is the primary function of interest, can be obtained as

$$\ln[I_l(x,y)/I_m(x,y)] = \mu ct(x,y) \qquad (3)$$

Hence, I_l and I_m must undergo logarithmic compression before simple subtraction is performed.

This technique of subtraction for removing unwanted background information from X-ray images was established about 40 years ago. It was conventionally implemented with photographic methods, utilizing the linear portion of the characteristic curve of X-ray film and the superposition of negative and positive transparencies [10]. Digital subtraction achieves the same objectives as photographic subtraction but offers far more advantages.

1.2.2 The Digital Approach

With the development of Digital Radiographic Systems, X-ray images are represented in computer readable form with each image broken into spatially discrete blocks, or pixels. According to the transmitted x-ray intensity, each pixel is assigned a corresponding grey tone value. Subtraction is then performed, pixelwise, between the mask and the live images. While temporal subtraction isolates the low contrast iodinated signal from the background structure, contrast enhancement is achieved by applying a post-subtraction gain to the image to fill the dynamic range (ie. the brightness range) of the display system. As opposed to photographic processing, digital subtraction is not limited by the linear region of the film charateristics curve, and is thus able to make use of a greater x-ray exposure range. Moreover, with the technical advance in x-ray imaging and processing systems, subtraction can be made in real-time and with improved ability to detect low contrast changes.

By using a digital system, the angiogram procedure is fairly simple. A small intravenous catheter is inserted into either the basilic or the cephalic vein (an arm vein), and then positioned near the superior vena cava (a major vein going to the right heart) to achieve higher arterial iodine concentration. The position of the catheter is checked fluoroscopically to ensure that it is in the correct position. The appropriate anatomic area of interest is then centered for imaging, and the injection is made. The volume of contrast agent to be injected varies with the anatomic area of interest and the site of injection. Typically, 30 to 50 cc are injected at a rate of 10 to 20 cc/sec using a power injector [34]. Since iodinated contrast agents are toxic and present some risk of serious complication [34], reducing the amount needed for the procedure would be an advantage.

The data processing sequence varies from system to system. In most cases, an image before the arrival of dye is acquired (either electronically or manually) as the mask. Subtraction is then performed between this mask and the subsequent live video images. The difference images, which enhance the iodine signal, are displayed on the video monitor during the actual processing and stored concurrently using an auxillary storage device. After the maximum concentration of contrast has passed through the area of interest, the process is terminated. The stored difference images may

then be reviewed and further processed or analyzed if necessary.

Intravenous injection, which was once abandoned because of its low contrast sensitivity, is once again gaining attention through the use of digital subtraction. The ability to enhance low contrast signals, coupled with the convenience and flexibility of digital systems make digital subtraction clinically attractive. With Digital Radiogrphy, intravenous arteriography has shown great promise as a means of assisting radiologists in patient diagnosis. The technique could lead to a major change in the way arterigraphy is praticed, and have a significant impact on the medical management of vascular disease. Various factors contribute to the success and can be tabulated as shown in Table 1.

TABLE 1

Potential of Digital Subtractive Angiography

HOSPITAL FACILITY ULTILIZATION	DSA offers low risk, out-patient based examination with reduced examination and physician time, thus increasing the efficiency of hospital facility ultilization. No x-ray film is consumed.
DIAGNOSTIC INFORMATION	The necessary information is obtained, in real-time, at a reduced amount of dose and with a less invasive process.
SIGNAL PROCESSING	The use of digital processing provides considerable flexibility in the method of subtraction and can be more accurate and sensitive than the conventional film based system.

1.3 ELEMENTS OF THE IMAGING SYSTEM

Because the subtraction is performed between two very large numbers to extract the small blood vessel signal, the imaging system must have low noise and wide dynamic range to record this small difference accurately. In its most basic form, the imaging system for DSA consists of an X-ray source, X-ray image intensifier, video camera, electronic amplifier, analog to digital converter and image storage device (Figure 2). The image intensifier converts the transmitted X-ray to light flux. The video camera then converts the light signal into an electronic (video) signal, which is then amplified. The amplification may occur before or after the digitization of data, and may be linear or logarithmic. After amplification, the images must be converted to digital form and processed. When processing is done, the result is ready to be displayed, and may be stored either in analog or in digital form. The relative merits of some of these options will be discussed below.

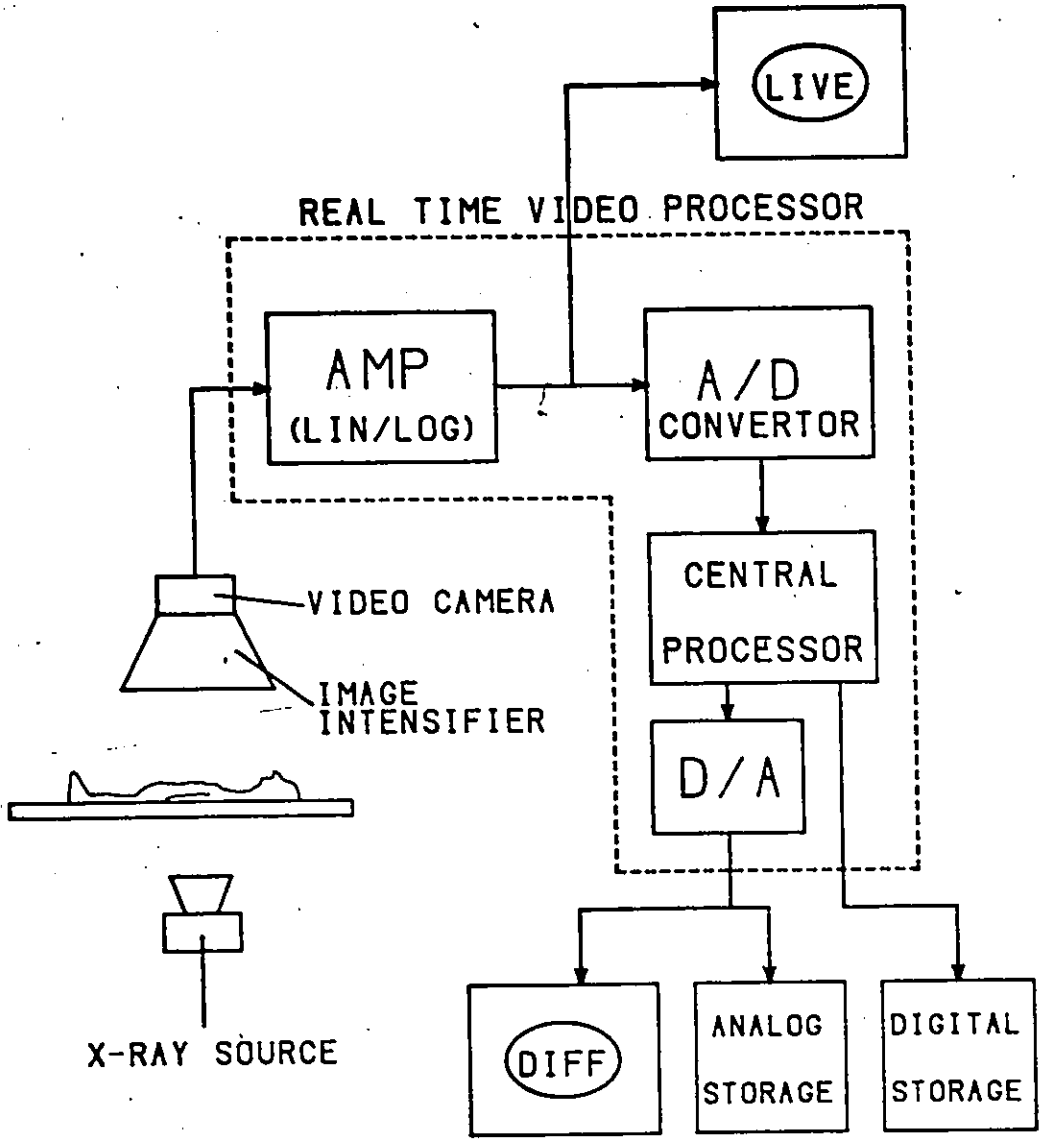


Figure 2: Block diagram and flow of information in a basic DSA system. The main imaging components include: x-ray source, an image intensifier, a video camera, A/D converter and image storage devices. The processed images are displayed on the TV monitor for on-line diagnosis.

1.3.1 X-ray source

Medical x-rays are produced by bombarding a tungsten target with high energy electrons and have wavelengths ranging from 10 to 50 pm. The low energy x-ray photons which are less able to penertrate the body, would only increase the patient dose and are thus removed by filtering. The remaining higher energy x-rays are collimated and then directed towards the patient. As the radiation passes through the patient, most is absorbed and only 1 to 4% is transmitted to the detector [68].

Due to the random nature of photon emission, the number of photons incident on any particular area 'a' during a fixed time period 't' fluctuates and follows a Poisson distribution. One generally defines the signal from the mean number $\bar{N}_x$ of x-ray photons incident on the area 'a' and defines the quantum noise from its standard deviation according to the Poisson statistics. The signal to noise ratio (SNR) is therefore

$$SNR = \frac{\bar{N}_x}{\sqrt{\bar{N}_x}} = \sqrt{\bar{N}_x} \qquad (4)$$

In order to achieve optimum performance and not waste x-ray dose, quantum noise should dominate the system noise of the imaging system, such that the SNR of the final image will not be further degraded.

1.3.2 Image Intensifier

The X-ray image intensifier performs the function of X-ray detector and light amplifier in a unit (Figure 3). X-rays entering the detector tube stimulate the input phosphor which emits light photons, and each photon can stimulate at most one electron from the photo-cathode. The electrons are then accelerated through a strong electric field and focused onto the anode. When the electrons strike the fluoresent anode, they excite and ionize the atoms of the material. Light photons are released when these excited atoms return to their usual state. As a result, a smaller but brighter image is produced with the brightness gain equal to the product of the geometric gain (ratio of the area of the input and output phosphors) and the electronic gain (the product of input quantum efficiency, photocathode efficiency, kinetic energy gain and the output phosphor efficiency).

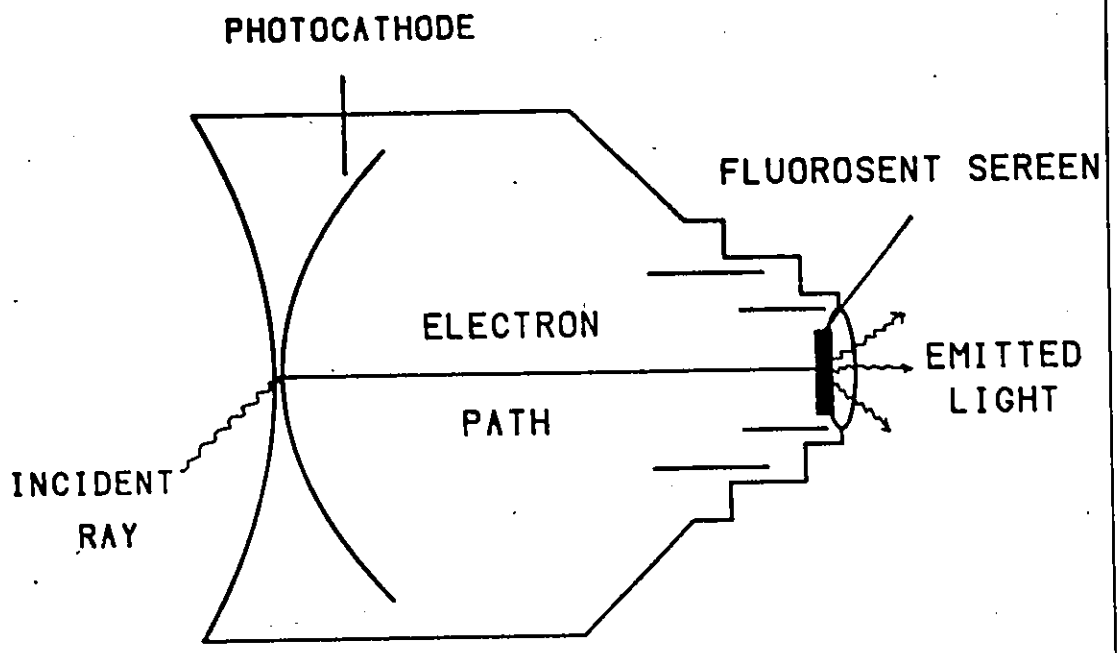


Figure 3: Image Intensifier exploits the photoelectronic effect, whereby the energy of incident photons is converted into the energy of moving electrons. Because electrons, unlike photons, are electrically charged particles, their energy can be increased by acceleration in an electric field. When photons strike a photosensitive metal plate, electrons are ejected into an evacuated region. The electrons are accelerated toward a postively charged fluoroscent plate, where their energy is converted again into energy of visible light.

1.3.3 Video Camera

The camera converts the light flux on the output phosphor of the image intensifier into an electronic signal which will eventually be digitized. Since the video camera system

is a major source of noise, many researchers believe that this is the most crucial element in the image chain. It should add little noise to the image, and a high performance TV link is necessary.

A specific example with typical values illustrates that a high SNR, state of the art video camera is required in order to be dose efficient [37,34]. With standard fluoroscopy where the dose[1] per frame requirement is 1μ R/frame, a camera with a SNR of 100:1 is quite appropriate (Figure[2] 4). However, in order to reduce quantum noise to the desired level for DSA, an exposure rate of 1 mR/frame is needed and dedicated video camera (SNR in the order of 1000:1) is desirable to keep the intrinsic system noise lower than the x-ray quantum noise [34].

---

[1] The unit of radiation is Roentgen (R) which is the amount of X-radiation that will produce 2.08 x 10' ion pairs per cubic centimeter of air at standard temperature and pressure.

[2] A thorough in-depth analysis of image noise is not the purpose of this chapter, a brief study is included in Appendix A. Interested reader should refer to Appendix A for a derivation of the image noise statistics which contributes to the plot of Figure 4.

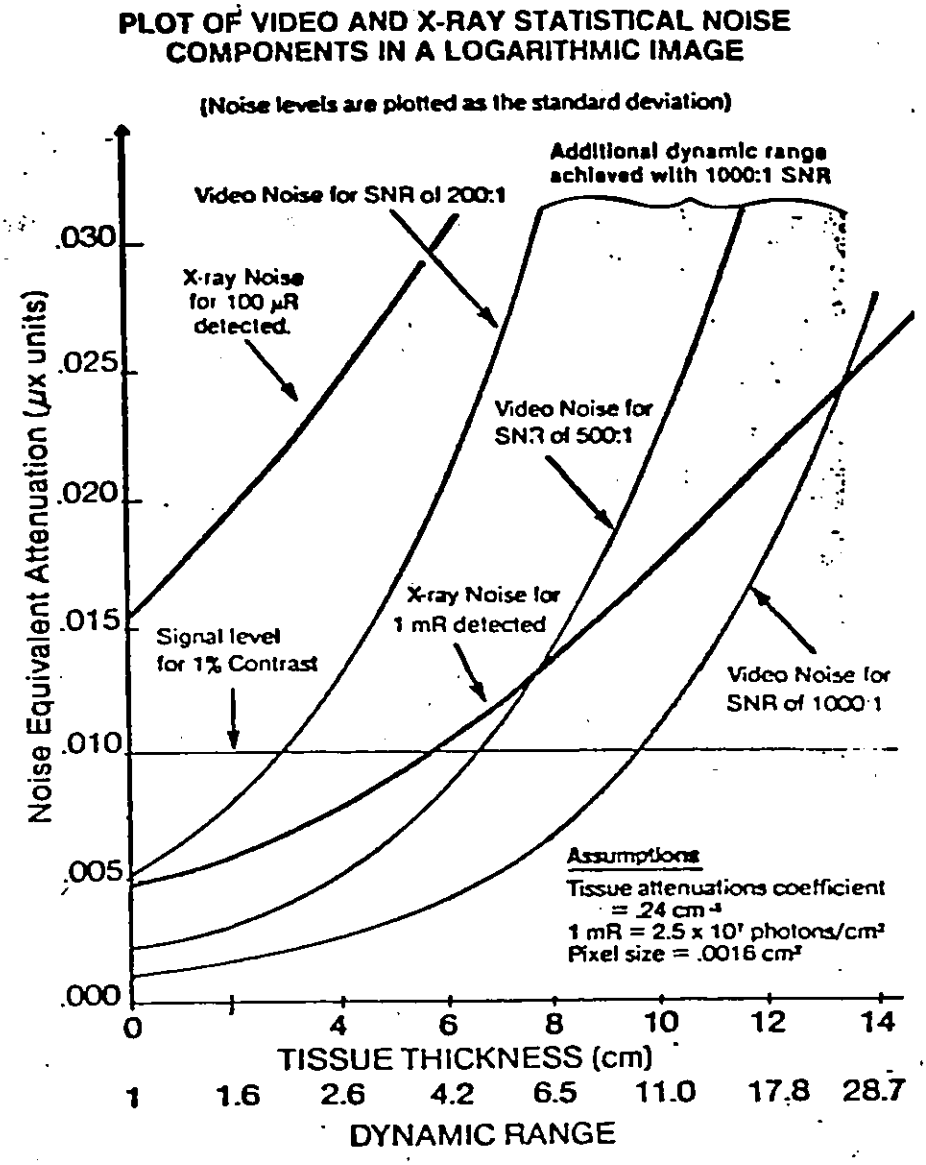


Figure 4: Noise components in a digital Radiographic image. A number of points are illustrated: 1) x-ray exposure must be high for the detection of small contrasts. 2) video SNR must be high to ensure that image noise is dominated by x-ray statistics. 3) with increased object thickness, the video noise component increases more rapidly than the x-ray component. (From Reference 34)

1.3.4 Amplifier

As derived earlier, logarithmic processing is required for the complete removal of the background structure. It can be shown that without such compression, an artery traversing regions of varying thickness appears in the subtracted image with non-uniform density [34]. In some systems, this constraint is relaxed, and the choice of amplification can be linear or logarithmic. While linear amplification is implemented directly, logarithmic amplification can be implemented either in an analog fashion using a logarithmic amplifier or in the digital form using a look-up table technique.

1.3.5 Analog to Digital converter (A/D converter)

The analog data, which has been broken into an array of spatially discrete chunks called pixels, must be converted to digital form for further processing. This conversion step is performed by an A/D converter with performance defined by the rate of digitization and the number of quantization levels. It must operate at sufficient speed to accommodate real-time video signal digitization (one sample every 70 ns for the North American NTSC system). The number of grey levels is determined by the dynamic range and the noise level of the analog signal, and the step size must be small enough so that the uncertainty associated with the di-

gitization process (quantization noise), does not add significantly to the overall image noise.

As an illustration, with a dynamic range[3] of 12 and a minimum noise (quantum + video) standard deviation of 0.007, (corresponding to the 1 mR exposure and a video SNR of 1000:1 in Figure 4), the number of grey levels required can be computed [37] by

$$\text{No. levels} = \frac{\ln 12}{0.007 \times 2} = 178 \qquad (5)$$

where the numerator represents the maximum logarithmic difference which must be accommodated by the digitizer and the denominator is the grey level spacing represented by the rms measure of the image noise after digitization. An 8-bit (256 levels) A/D converter is therefore adequate for this example.

1.3.6 Image Storage

After the digital images are processed, they are stored in auxillary memory devices to be ready for further examination. Because digital data is the basis for all image manipulation, it would be an advantage to store all data in digital form. Due to the high cost and low operating speed

[3] This is computed by the ratio of the maximum video signal to the minimum video signal and can be written as

$$R = S/Se^{-\beta} = e^{\beta}$$

where β is some attenuation coefficient

of available digital storage devices, the digital nature of the data is preserved usually with a decrease in the amount of information present (eg. a decrease in spatial resolution or frame rate) in an angiogram study. With enough high speed digital storage devices in parallel, the required speed can be achieved but at a very high cost. Because of these factors, analog devices continue to be commonly used for the real-time storage of high frame rate, high resolution image sequences.

Analog storage devices can operate at real-time video rate, but typically have a maximum SNR [34] of 100:1 to 200:1. This would seem unusually low compared to the TV camera SNR requirement of 1000:1. The key can be to store only the amplified difference images. If the difference images are stored with the iodine signal amplified to occupy a large portion of the video dynamic range, the resulting SNR is effectively increased by a factor equal to the brightness gain which is applied to the difference images prior to storage (Figure 5). Typically, the difference can be amplified by a gain of 8 and still not exceed the dynamic range [58].

It may be of advantage to save all unsubtracted raw images and post-process them to achieve low contrast enhancement, at a later time. Post-processing of data has the advantage that sophisticatd image processing and analysis

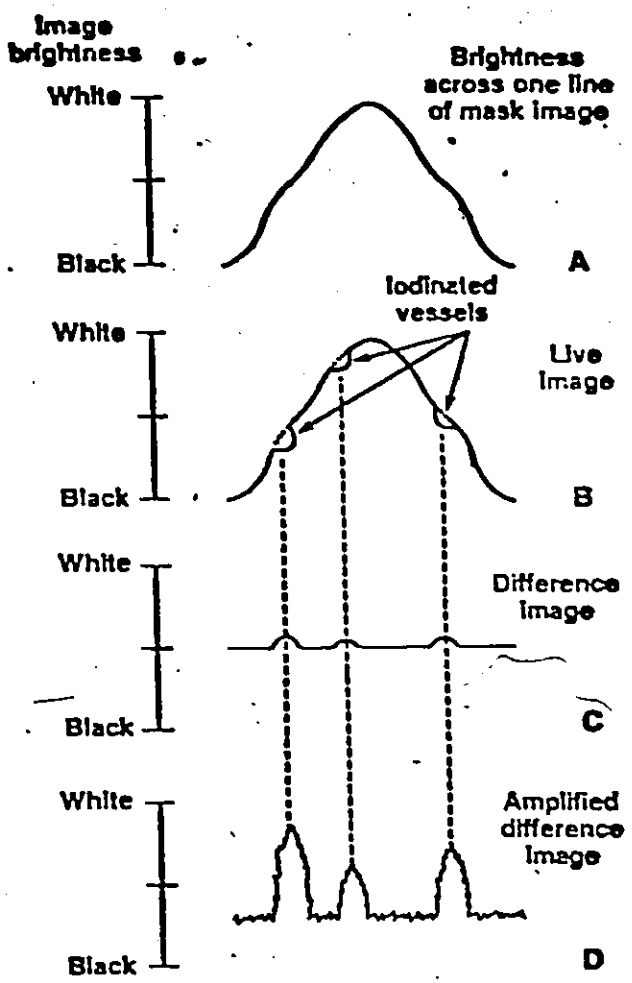


Figure 5: A temporal subtraction sequence showing the effect of amplifing a difference image. The difference obtained from a simple subtraction is often imperceptible. Amplifing the difference to fill the brightness range of the display helps to improve the vessel contrast, but at the expense of increasing the background noise. (From reference 34)

techniques which are not feasible for real-time implementation can be applied at a later time after the data is acquired and the patient procedure is completed. However, the

relatively low SNR of analog storage devices and the limited space of digital storage do not make this feasible.

1.4 LIMITATIONS: IN DSA

1.4.1 Patient Motion

The primary limitation to successful intravenous DSA is the patient motion that occurs between the mask exposure and the contrast exposures. Very often, motion of less than 1 mm. of a high contrast interface (soft-tissue/bone) will produce significant distortion and artifact in the subtracted images [6] (Figure 6). The effects of motion are generally blurring of vessels, or streak artifacts consisting of light and dark bands at opposite edges of a high contrast object, such as bone. In general, the lower the concentration of iodine, the more susceptible the image will be to motion artifact. Motion artifacts are difficult, if not impossible, to correct. Once a run is started, nothing can be moved. The effects and remedies for patient motion are considered in detail in chapter 2.

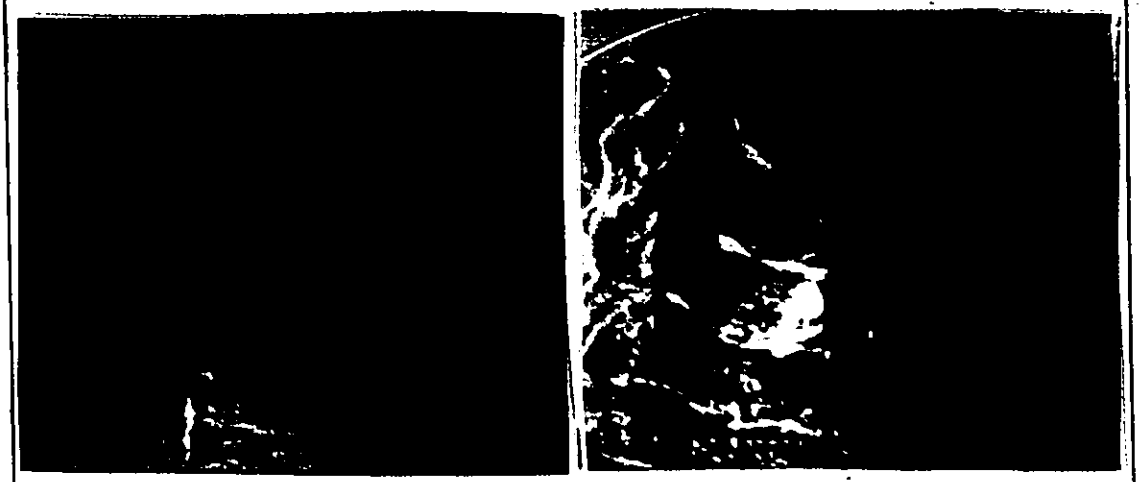


Figure 6: Effect of Patient motion in DSA. a) An intravenous arteriogram demonstrating low contrast enhancement of DSA. b) Temporal subtractive image demonstrating artifacts caused by swallowing. Artifacts as such destroy much of the diagnostic information available.

1.4.2 Noise

In DSA, the signal is determined by the amount of iodinated contrast agent in the blood vessel and by the average energy of the X-ray beam. Unlike conventional radiography where image contrast determines how well structures can be seen, the visibility in DSA is determined by the SNR of the subtracted images. The SNR in DSA is theoretically limited [6] by the iodine concentration [I] and the x-ray dose D according to the equation

$$SNR = k[I]\sqrt{D} \quad (6)$$

where k is a constant of proportionality determined by the noise level of the imaging system.

Although the equation suggests that it is more efficient to increase the SNR by increasing iodine concentration than x-ray dose, due to the toxicity of iodinated agent, radiation dose remains the only parameter used to improve vessel visibility. Furthermore, if the noise level is not properly controlled, the visibility will be significantly impaired.

The subtraction technique is fundamentally limited by the noise level of the subtracted images which can be catagorized into the structured and the stochastic components. The structured portion is mainly caused by image tube distortion and patient motion, which has been discussed. The stochastic portion consists of the dose dependent X-ray quantum noise and the additive electronic noise. The dose dependent noise is fixed by the number of photons passing through the patient (proportional to dose) and by the efficiency of the X-ray detector (percentage of incident photons actually absorbed by the detector). The additive noise component is the electronic noise introduced by the electronic system (eg. camera).

Of the four sources mentioned, only x-ray quantum noise is normially under the control of the operator. In order to

be dose effective, the dose dependent x-ray quantum noise should dominate the others, so that the signal to noise ratio can be improved by increasing the radiation exposure, and decreased by lowering the dose. With high additive and structured noise, the subtracted images will fail to demonstrate blood vessels. In such cases, an increase in X-ray dose, however, will not necessarily improve the quality of the subtraction.

Chapter II

PATIENT MOTION IN DSA

Although digital subtraction has been shown to be an effective technique for contrast enhancement in medicine, it is primarily limited by sudden patient motion. Due to patient discomfort caused by the iodinated contrast material and to the amount of time required for the injected contrast agent to travel to the artery under examination (>3 sec), patient motion artifacts are present, to a certain degree, in almost all digital intravenous angiograms. All too often, such artifacts destroy much of the diagnostic information available.

There have been several modifications to DSA since its implementation in 1980. Based on different processing schemes, each modification attempts to address the problems in DSA differently. Their relative abilities to handle patient motion is still an unknown, and a definitive study comparing conventional DSA with the various integration techniques on the basis of motion artrifact, is yet to be performed. [6,36,60]

2.1 MOTION ANALYSIS

All structures in the human body are subjected to temporal changes. Although the human visual system has an outstanding capability to recognize complex spatial patterns, automated recognization of complex motion patterns as they occur in an x-ray image sequence is often very difficult.

Classification schemes for patient motion have been proposed [27,36]. In general, the kinds of motion that usually occur during imaging can be classified as either oscillatory or non-oscillatory. Oscillatory motion includes cardiac pulsation, breathing and swallowing, and produce image artifacts that repetitively appear, disappear and reappear with reverse phase. Non-oscillatory motion is usually unpredictable and is the most troublesome to handle. Sudden patient motion falls into this second catagory and is caused mainly by the subtle movement of rigid objects such as the head and the limbs. Table 2 is a list of the temporal frequency ranges associated with the various temporal changes. Notice that physiological events that vary rapidly (eg. cardiac pulsation) have higher temporal frequencies than events that vary more slowly (eg. the arrival and washout of contrast material). It will be shown in later sections how physical phenomena and a priori knowledge as such may be used in the handling of patient motion.

TABLE 2

Temporal frequency range of Various Motion Type [36,27]

Breathing	10-20 cyc/min
Swallowing	<40 cyc/min
Cardiac Pulsation	>60 cyc/min
Contrast Opacification	a duration of 3-10 sec

2.1.1 Cardiac Pulsation

In arterial pulsation, the artery spends most of the time at or near its end-diastolic position with sharp periodic systolic impulses and at a rate of more than 60 cycles/min. It is associated with the expansion and contraction of the heart chamber. Because the pulsation is periodic and is at a rate higher than that associated with the rate of change of the iodine bolus, a satisfactory results can be obtained with a subtraction in which only images of the same cardiac phase are subtracted. This provides a better registration among images.

2.1.2 Breathing

The temporal variation associated with breathing most resembles that of iodine bolus. It ranges from 20 cycles/min. for fast breathing down to 10 cycles/min. for slow breath-

ing. Although patients are generally asked to hold their breath during imaging, for a severely ill patient this is not generally successful. A way of reducing the effect of breathing is through temporal filtering. Instead of subtraction, the image sequence may be bandpass filtered to remove the stationary background and the temporal variation of high rate. A single frame from the filtered image sequence may then be captured with less breathing artifact and yet showing good arterial opacification.

2.1.3 Swallowing

Swallowing occurs in about 30% of most intravenous carotid arteriograms [5]. It presents complicated dynamic variations and covers a wide domain of frequency range. Though the motion is not periodic, an image area corresponding to the larynx region undergoes oscillatory behavior during swallowing. As in the handling of breathing, a video frame with reduced artifact may be captured from the bandpass filtered image sequence. If such is not possible, swallowing artifact can also be removed by choosing a new mask, and hopefully some images of the sequence will be adequately registered.

2.1.4 Sudden Patient Motion

This refers to abrupt patient movement from position A to position B and occurs in approximately 10% of all angiogram examinations [34]. Because of patient discomfort caused by the iodinated contrast material, this type of motion is not uncommon. At a pixel level, its motion artifact can be modelled by a step function changing the pixel intensity from one level to another. The best way to handle this problem is, perhaps, by reselecting the mask image, as in the case of swallowing, so that the succeeding subtraction will be registered.

2.2 MOTION COMPENSATED TECHNIQUES

In order to render solutions to the limitations in DSA, a number of investigations [5,35,18] have been made to study various DSA related phenomena, such as bolus dynamics and alternative DSA image acquisition and processing schemes. Among these are motion compensation schemes including remasking, energy subtraction, temporal bandpass filtering and image matching.

2.2.1 Remasking

Motion artifacts caused by misregistration between the mask and the live images can be suppresssed by allowing the selection of an alternate mask that better matches the ori-

entation of the background anatomy. Clinical study [34] has shown that 30-60% of temporal subtaction studies can be improved by remasking. If live unsubtracted images are stored, this is a straightforward procedure. Subtraction is simply performed retrospectively betweem the contrast-filled live image and the alternate mask image, producing a difference with minimum motion artifact. In the case when only the difference images are stored, remasking can be done in a similar way. Let the original mask image be designated M, and the unsubtracted live images L1, L2,.....Ln. The original differences are then given by

$$D1 = L1 - M$$
$$D2 = L2 - M$$
$$\vdots$$
$$Dn = Ln - M \qquad (7)$$

Suppose D10 has the maximum iodine contrast, but has significant misregistration artifact due to motion. And suppose L5 has little iodine contrast but better matches the orientation of anatomic structure of L10. The desired image is, therefore, the difference between L10 and L5. With only D1,D2,...Dn available after image storage, the desired difference can be obtained if D5 is subtracted from D10, producing

$$D10 - D5 = (L10-M) - (L5-M)$$
$$= L10 - L5 \qquad (8)$$

As mentioed in the last chapter, this post-processing method requires a high SNR in both D10 and D5. Because the difference images are amplified before storing and because the x-ray statistics noise in the original mask M cancels out deterministically during the subtraction, no increase in x-ray statistics noise is introduced. However, if the dynamic range in either of the amplified difference images is exceeded in the region of motion, there is a danger that the iodinated vessel signal, which may be in the region of motion, may not be retained.

2.2.2 Energy Subtraction

Whereas, temporal subtraction depends on the change of iodine concentration with time, energy subtraction utilizes the difference in attenuation properties between contrast agent and the surrounding soft tissues or bones at different x-ray photon energies.

Energy subtraction is based on the fact that the x-ray attenuation of matter is energy dependent and that the variation is different for materials having different average atomic numbers (Figure 7).

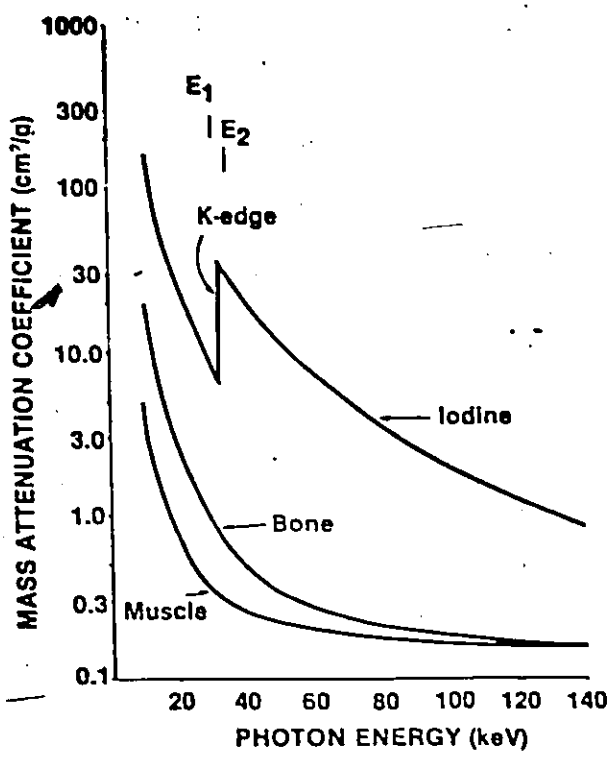


Figure 7: Energy dependence of the attenuation coefficients of iodine and anatomic structure. Images made at energies E1 and E2 are subtracted from each other. Because the attenuation coefficients of soft tissue and bone change only slightly for the two energies while that of iodine changes significantly, the relative ratio of contrast (iodine/soft-tissue) is increased by using the difference. (From Reference 13)

In this process, images are acquired at x-ray energies E1 and E2 slightly below and above the large discontinuity (K-edge) in the attenuation curve of iodine at 33 KeV. Because the high and the low energy images can be obtained within a few milliseconds interval, the subtraction is relatively insensitive to motion. However, the imaging system for energy

subtraction is more complex than that required for temporal subtraction. A second drawback to the approach is that the difference image obtained generally contains non-zero background anatomy. In other words, not all of the background anatomy can be removed. In tissue cancellation, there are residual signals from bones present in the object and vice versa.

Suppose the object to be imaged consists of only iodine and a non-iodinated soft-tissue material. And further, suppose that each x-ray beam is monoenergetic, the transmitted intensity of the two beams would be

$$I_1(x,y)=I_{10}\ \exp\{-[\mu_{t_1}Z_t(x,y)+\mu_{i_1}Z_i(x,y)]\}$$

and

$$I_2(x,y)=I_{20}\ \exp\{-[\mu_{t_2}Z_t(x,y)+\mu_{i_2}Z_i(x,y)]\} \qquad (9)$$

where μ_i and μ_t are the attenuation coefficients of each material,

Z_i and Z_t are the projected thickness,

I_1, I_2 and I_{10}, I_{20} are the transmitted and incident rays of the high and low energy beams respectively

By defining

$$L_1 = \ln(I_{10}/I_1) = \mu_{t_1}Z_t + \mu_{i_1}Z_i$$
$$L_2 = \ln(I_{20}/I_2) = \mu_{t_2}Z_t + \mu_{i_2}Z_i \qquad (10)$$

a weighted subtraction which enhances the iodine contrast can be obtained as

$$\begin{aligned} D &= K_1 L_1 - K_2 L_2 = \mu_{t_2} L_1 - \mu_{t_1} L_2 \\ &= (\mu_{t_2}\mu_{i_1} - \mu_{t_1}\mu_{i_2}) Z_i + (\mu_{t_2}\mu_{t_1} - \mu_{t_1}\mu_{t_2}) Z_t \\ &= (\mu_{t_2}\mu_{i_1} - \mu_{t_1}\mu_{i_2}) Z_i \end{aligned} \tag{11}$$

where only the iodine contrast contributes.

In general, to cancel out contrast from an arbitrary material, the weight factors k_1, k_2 must satisfy

$$\frac{k_1}{k_2} = \frac{\mu_{m_2}}{\mu_{m_1}} \tag{12}$$

where subscript m refers to the material to be cancelled.

The assumption that only one noniodinated material is present is not true in general. Therefore, with bone and soft tissue both present, only one material can be cancelled, leaving all other materials to cause residual signals.

2.2.3 Temporal Bandpass Filtering

The bandpass filtering technique is based on the theory that the time varing contrast flow signal can be isolated through a bandpass filter if the passband of the filter overlaps, as much as possible, the temporal frequency components associated with the contrast bolus. The response to

signals with frequencies outside the filter passband will therefore be minimized.

If the filter's input and output signal are denoted as s(i) and s'(i), respectively, and the impulse response of the filter is denoted as f(i), then the output s'(i) can be obtained as the convolution of s(i) and f(i)

$$s'(i) = \sum_{j=i-M}^{i} s(j)f(i-j) = s(i)*f(i) \qquad (13)$$

with f(i) defined in the integer range $0<i<M$ and s(i) defined in the integer range $0<i<N$. (Note that * denotes convolution, N is the number of samples and M is the filter length)

Denoting f(i-j) as k_j, s'(i) is simply a weighted sum of all the M previous input samples ending with the i-th sample. Equation 13 therefore becomes

$$s'(i) = \sum_{j=i-M}^{i} k_j s(j) \qquad (14)$$

As a rule of thumb, to ensure that the stationary background anatomy in the i-th output is cancelled, an additional constraint is added so that

$$\sum_j k_j = 0 \qquad (15)$$

One can easily see that conventional temporal subtraction is, in fact, a special case where

$$k_j = \begin{cases} +1 & ,j=1 \\ -1 & ,j=m \\ 0 & ,\text{all other } j \text{ values} \end{cases} \qquad (16)$$

with 1 corresponding to the time of maximum opacification and m to a time of low opacification.

The technique can also be implemented using a recursive digital filter. With reference to filter 1 in Figure 8, the filtered image at time t_o is a weighted sum of many preceeding video frames with decreasing emphasis further in the past frames. Filter 1 principally adds together video frames of the previous 2 seconds. Filter 2 has a much longer decays but with the integrated area under it equal to that of filter 1. The actual resulting image observed at time t_o, is then the difference between the output of filters 1 and 2. Since the area under filters 1 and 2 are equal, the sum of all weights is zero. This ensures complete background subtraction.

The implementation is made using the formulas given by

$$s_1'(j)=n_1 s_1'(j-1)+(1-n_1)s(j-1)$$
$$s_2'(j)=n_2 s_2'(j-1)+(1-n_2)s(j-1) \qquad (17)$$

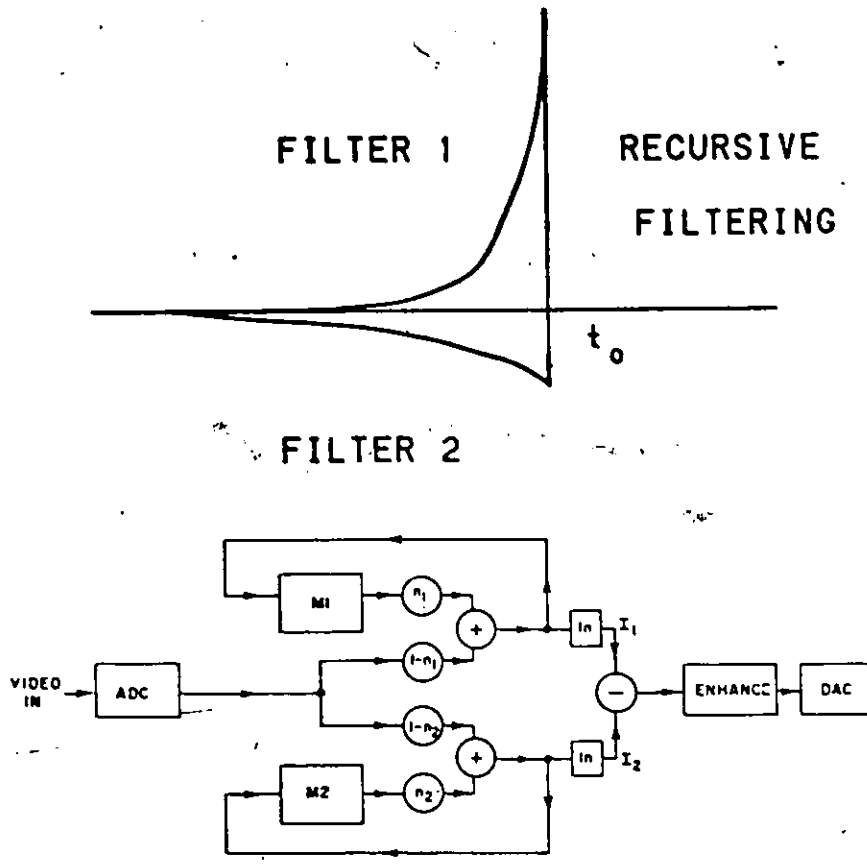


Figure 8: Recursive filtering. a) weighting coefficients: The resulting image is obtained by multiplying each original image by the respective weight and adding all weighted image together. b) block diagram of an implementation scheme employing two separate memory units. (From Reference 38)

where $s_i'(j)$ is the measured output of memory M_i at time j and s(j) is the new input pixel value at time j.

The quantity $s_1'(j)-s_2'(j)=s'(j)$ is the bandpass filtered output. The frequency response H(z) for this filter is given as

$$H(z) = \frac{(z-1)\ (n_1 - n_2)}{(z-n_1)(z-n_2)} \qquad (18)$$

H(z) has one zero at z=1 (w=0) and two poles, one at $z=n_1$ and the other at $z=n_2$. The impulse response is therefore given by

$$h(t) = \eta_1 e^{-\eta_1 t} - \eta_2 e^{-\eta_2 t} \qquad (19)$$

which starts at zero, rises to a maximum and decay with a long exponential tail. (note: η_1 and η_2 are the time constants of filter 1 and filter 2 respectively) This is similar to the contrast flow curve.

In recursive bandpass filtering, the filter's time constants η_1 and η_2 are selected before the DSA run. In other words, the filter is prospective. Therefore, only if the frequency range of motion lies outside of the passband of the predetermined filter characteristics will motion effects be removed. Trade-offs between having a narrow passband for motion removal and having a wide passband for contrast enhancement thus exist.

2.2.4 Image Matching

Correction for rigid translational motion between images can be accomplished by maximizing the two dimensional cross-correlation function and repositioning the images before

subtraction. In fact, this has been a standard image processing technique. Assuming that the misalignment is caused by rigid translational motion, the two dimensional translational shift between the two images are computed by image matching. Correction is then made by subtracting the corresponding pixels of the two images after realignment.

Let the mask image be given by a matrix $m(i,j)$ of grey values with integer subscripts. From m, a normalized matrix m' is computed by subtracting from each value the mean grey value of m and then dividing by the standard deviation of the grey value of m. Next, the normalized matrix m' is used to define a function $m'(x,y)$ for real number x and y through interpolation. The cross-correlation function between the live image $c(i,j)$ and the mask $m'(x,y)$ is then

$$C(\mu,\nu) = \sum_i \sum_j c(i,j)m'(i+\mu, j+\nu) \tag{20}$$

where the summation is carried out for a small image region or block.

The location of the maximum correlation coeficient then provides a measure of the amount of the translational misalignment. The images can then be realigned by unshifting the mask image by this amount prior to subtraction.

Unless a priori knowledge of the estimated shift is available, the above correlation process requires hours of

CPU time even for a dedicated minicomputer, and is therefore not appropriate for real time purposes.

2.3 OTHER MODIFICATIONS TO CONVENTIONAL DSA

There are many other modifications to the conventional DSA system. They are, in one way or another, transformations or combinations of the above versions, and their abilities to remove motion artifact are sometimes uncertain. Some important techniques include matched filtering, vascular tracing and hybrid subtraction.

1. MATCHED FILTERING

   The matched filtering technique is based on the principle that the highest SNR possible is attained if the filter response is matched to the shape of the signal of interest. The matched filter response to a sudden patient motion, therefore, will be less than to a contrast bolus (Figure 9). However, the process relies on the ability to obtain the filter weight coefficients from the contrast dilution curve. This can be done by placing a region-of-interest cursor on the image region of the artery and measuring the local contrast dilution curve. However, because the contrast dilution curve varies from place to place and from artery to artery, the process has to be performed locally. Furthermore, misalignment of the image due to motion will cause a problem in obtaining

the contrast dilution curve and hence the filter coefficents.

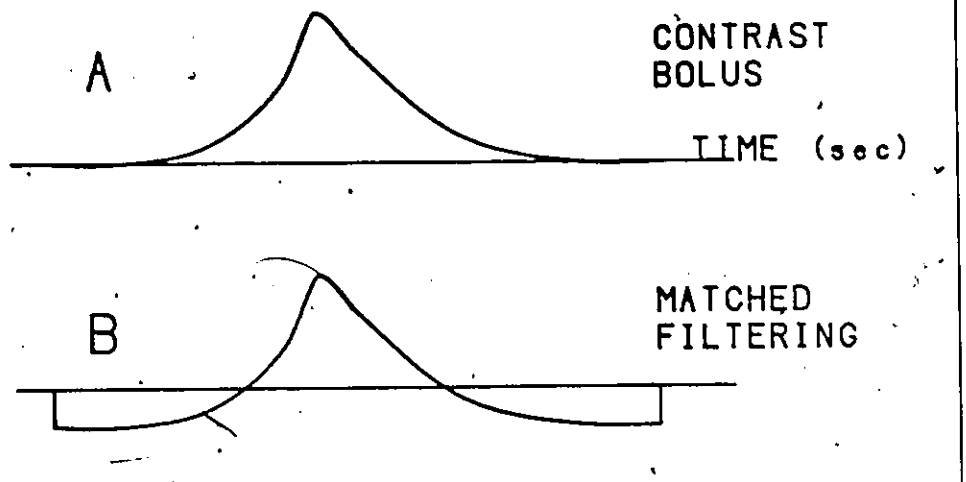


Figure 9: Matched filtering. a) a plot of typical contrast bolus flow curve following intravenous injection. b) a plot of the weighting coefficients for matched filtering. Because the weights are made to match with the bolus flow curve, the filter response is maximum for the contrast bolus and lower for undesired patient motion.

2.1 VASCULAR TRACING [40]

Because contrast bolus is of a short duration, only a small segment of vasular anatomy would normally be shown at any given moment. The vascular tracing technique, however, records maximum opacification of vascular structures as they occur and leaves them stored, pixelwise, in the memory of the video proces-

sor. Therefore, after the contrast bolus has passed through all imaging structures, a single image is formed which displays all the vascular anatomy with maximum opacification, independent of when maximum opacification occurs. It basically stores, for each pixel, the most negative difference over time. However, the technique fails when any slight motion occurs, In the case of motion, the technique would store the dark streak caused by motion instead of the maximum iodine opacification.

3. HYBRID SUBTRACTION

This is a technique involving both the temporal and energy subtraction techniques (Figure 10). It combines dual-energy subtraction for the removal of soft tissue structures and temporal subtraction for the elimination of residual bone structure. The pre- and post-contrast dual energy image pairs are obtained with each high and low energy pair subtracted to remove soft tissue structure. Then, the resulting tissue cancelled pre- and post- contrast images are subtracted to remove the residual bone structure. As a result, soft tissue motion can be eliminated. However, when both soft-tissue and bone are involved in the motion process, only one type of motion can be removed.

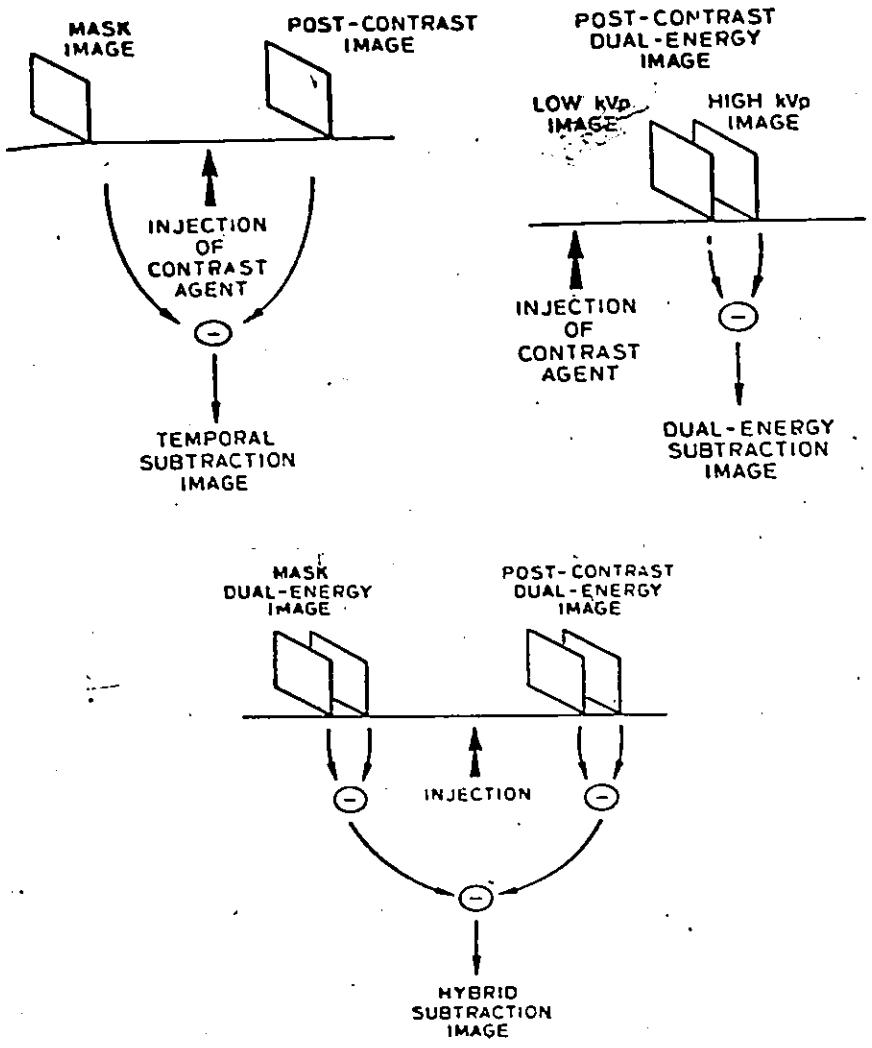


Figure 10: Relationship of hybrid subtraction to temporal and energy subtraction. a) Temporal subtraction. b) Energy subtraction producing a soft-tissue-cancelled image. c) Hybrid subtraction: Pre- and Post- contrast dual-energy image pairs are obtained. First, each pair is subtracted to remove soft-tissue variations. Then the resultant tissue-cancelled pre-and post contrast images are subtracted to remove the residual bone structure. (From Reference 5)

2.4 SUMMARY

Patient motion is the primary limitation to intravenous angiography. But because of its complexity, it is difficult to describe and handle analytically. Unless one is able to stop the physical motion of the patient, the motion problem cannot be ignored.

Various methods showing improvement over direct DSA have been investigated during the past few years. Remasking, energy subtraction, image matching and bandpass filtering each seems to respond to the motion problem in a different fashion. The bandpass filtering approach tends to remove periodic motion of high frequency (>30 cyc/min). Energy subtraction is immune to patient motion, but leaves bone structure as residual signal in the difference image. Remasking and image matching techniques seem to address themselves to sudden patient motion but can be time consuming. Each technique has its own advantage and drawback. Moreover, clinical considerations such as safety, cost, facility ultilization and speed must also be consideration.

Chapter III

REAL TIME REMASKING AND MOTION ESTIMATION

Due to all the difficulties that motion imposes, there is no easy way to the development of automated systems which will handle the motion problem efficiently. Presently, in order to ensure that clinically valuable data is obtained in an examination, the angiogram procedure is repeated whenever the effect of motion is thought to be significant. In the process, both patient and physician will be exposed to radiation which would not have been necessary if patient motion did not occur.

Although some motion compensated techniques for motion correction do exist, very little clinical pratice has been gained to date because of the lack of technical personnel and the time required for the post-processing of the image sequence [6,62].

It would, therefore, be an advantage to have a real-time automated technique for removing motion artifacts thus freeing both patient and physician from exposure to excessive radiation by reducing the need to repeat procedures.

3.1 REAL-TIME REMASKING (A PROPOSED NEW TECHNIQUE)

Because of the time it takes for the contrast bolus to travel to the artery of interest, there is usually a few seconds delay in the arrival of contrast bolus after the injection. As a result, patient motion occuring during this time interval is common. Therefore, it is theoretically benefical to have the mask frame acquired as close to the contrast arrival as possible, in order to reduce the possibility that the subtraction be affected by sudden patient motion.

Real time automated remasking is a technique proposed in this dissertation to achieve this goal. The technique attempts to implement the previously studied remasking scheme in real time during data acquisition. As a result, motion correction is done as data is being collected and neither off-line post-processing nor re-exposure is needed. In the process, a simple real-time motion detector is incorporated into the DSA system (Figure 11). During the process of data acquisition, a new mask is acquired whenever significant patient motion is detected. Thus the image misregistration in the succeeding subtraction is minimized.

As an illustration, consider the different cases shown in Figure 12 . While patient motion results in motion artifacts and a decrease in iodine contrast in the original subtraction, real-time selection of a new mask leaves the iod-

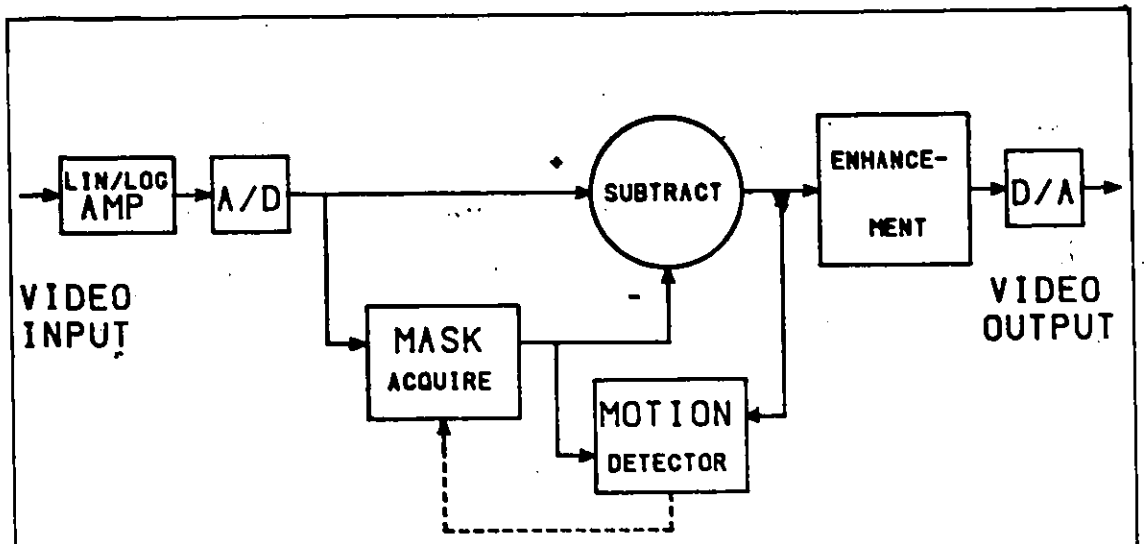


Figure 11: Block diagram of a real-time remasking system. The motion detector analyses the mask and the difference images. A decision is then made whether to update the mask. The control line is shown as a dash line in the diagram.

ine contrast signal unaffected. More importantly, the high contrast, non-zero background that is present in the original difference image is removed through this automated remasking technique. Misregistration is ,therefore, completely avoided if motion occurs well before or after the dye arrives. If, however, motion occurs during the arrival of dye, a decrease in dye contrast may result due to the fact that the new mask may contain some contrast agent. Fortunately, the flow of dye appears as a wavefront washing out of one segment of the vessel and moving on to the next. Thus, only a small section of the vessel is affected, while

misalignment is being corrected for the succeeding subtractions.

However, this real time automated remasking approach is not appropriate for the handling of cardiac pulsation motion. With the process described, regular motion as such would lead to frequent changes of the mask and would thus significantly reduce the iodine contrast signal. Satisfactory results can therefore only be obtained when improvement is expected with the post-processed remasking technique described in the previous chapter.

In addition, because the success of the automated remasking system relies very much on the correct detection of motion, the detection or estimation of motion is a critical function in the process. In general, if a particular difference image resulting from the subtraction of two angiogram images is considered, three possible hypotheses may be concluded: Either the two images are the same and differ only due to random noise, the images are different due to the arrival of contrast agent or the images are different due to motion. Since the arrival of dye is a desired signal while patient motion is not, the detection criterion must be able to discriminate significant motion from dye arrival and must be simple enough for real-time implementation. The detection problem in the context of DSA will be discussed in chapter four. In the remainder of this chapter, the problem

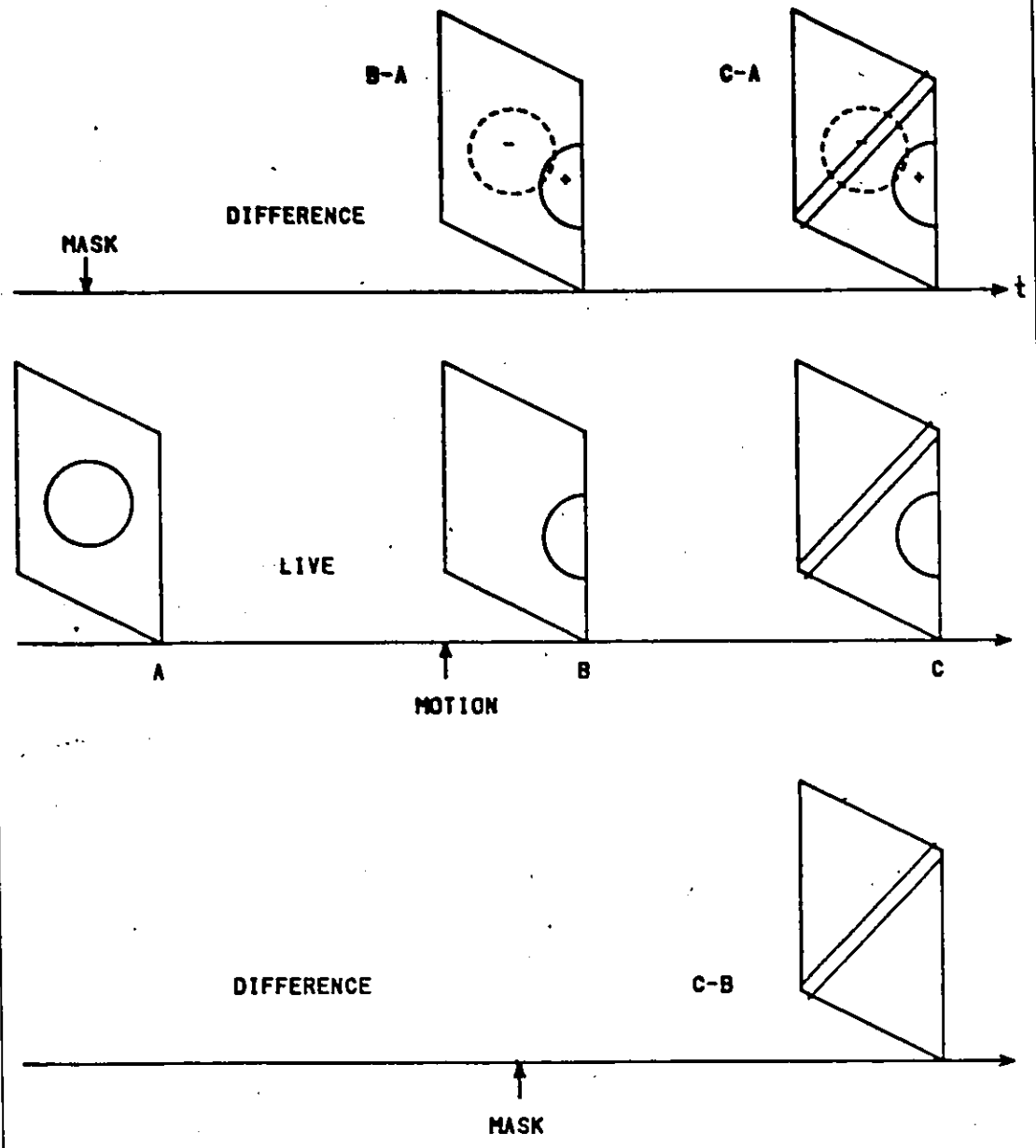


Figure 12: The effect of real-time remasking. With A as the mask, (C-A) shows vessel opacification associated with motion artifacts. With B as the mask, (C-B) shows a clean vessel opacification signal.

of motion estimation is studied from a general image processing point of view. The same principle applies in the angiogram case.

3.2 ESTIMATION OF IMAGE TRANSLATION

Translational motion is often assumed in many image processing applications involving motion. Strictly speaking, it is not correct to describe real-world motion with two dimensional translation motion, but the assumption can be considered valid when the sequence of images are acquired at a high rate, so that the relative motion between any two consecutive image frames is essentially translational. In the DSA application, the translational model of motion is perhaps the only practical way to the analytical study of real-time remasking and it shall be assummed that patient motion can be adequately described by the translational motion. The three major approaches to the estimation of two dimensional image translation are: Fourier method, image matching and the method of differentials.

3.2.1 Fourier Method

In the Fourier transform domain, translation due to a shift in the spatial coordinate is known to introduce a corresponding linear, frequency dependent phase shift in the frequency domain. With (x,y) denoting the spatial coordi-

nate, (u,v) the spatial frequency and F(u,v) the Fourier transform of f(x,y); the Fourier transform G(u,v) of g(x,y), a translation of f(x,y) is given by

$$G(u,v)=F(u,v)\exp[-2\pi j(u\Delta x+v\Delta y)$$

with

$$g(x,y)=f(x-\Delta x,y-\Delta y) \tag{21}$$

Furthermore, the difference between the phase angles of F(u,v) and G(u,v) is obtained as

$$\theta(u,v)=\angle G(u,v)-\angle F(u,v)=-2\pi(u\Delta x+v\Delta y) \tag{22}$$

Therefore, by calculating θ at two frequency pairs (u,v), it is possible to solve for Δx and Δy.

However, the phase angles of F(u,v) and G(u,v) are not unique. In order to prevent any ambiguity, one can add the constraint

$$|\theta|<\pi \tag{23}$$

With the assumption that the amount of translation $(\Delta x,\Delta y)$ is less than half of the frame size (that is L/2), one can obtain a unique estimate of $(\Delta x,\Delta y)$ by calculating θ at frequencies $u<1/2L$, $v<1/2L$.

A major limitation to the Fourier method is when the motion is not that of a simple image translation but of an ob-

ject moving in a stationary non-uniform background. In such a case, g(x,y) is no longer a shifted version of f(x,y), and the Fourier method breaks down. Therefore, in the estimation of object motion with a stationary background, it is required to first segment the images into stationary and non-stationary areas. This segmentation requirement is, in fact, common in applications such as coding and computer vision. Because of these complications, the Fourier method is not widely used.

On the other hand, with the convolution theorem of Fourier analysis, the Fourier transform technique can be used in conjunction with the correlaton method, a technique to be described in the next section.

3.2.2 Image Matching

In the matching method, a search procedure is implemented to find the relative translation between images. By shifting a small region (block) of an image, f(i,j), over a large temporally different image g(i,j), the translational shift is located at the point where a maximum of a similarity measure is found. In mathematical form, this can be shown as

$$S(\Delta x, \Delta y) = C\{f(x,y), g(x,y)\} \qquad (24)$$

where

C is a cost function

$g(x,y)=f(x-\Delta x, y-\Delta y)$

and $f(x,y)$ represents a small image block (window)

Perhaps the most commonly used cost function is cross-correlation. It is based on the statistical property that the autocorrelation function, $R(\tau)$, of a stationary process $x(t)$ is maximum at $\tau=0$ [55]. For a stationary process $x(t)$, its autocorrrelation function is given by

$$R(\tau) = E\{x(t+\tau)x(t)\} \tag{25}$$

It can easily be seen ,by squaring the function, that

$$E\{x(t+\tau)\pm x(t)\}^2 = 2[R(0)\pm R(\tau)] \tag{26}$$

Since the left hand side is obviously non-negative, hence

$$R(0)\pm R(\tau)\geq 0$$

that is

$$R(0)>R(\tau)\geq -R(\tau) \tag{27}$$

Thus $R(\tau)$ is a maximum at the origin

$$|R(\tau)|\leq R(0) \quad \text{for all} \tag{28}$$

In terms of a discrete two dimensional image domain (Figure 13), the cross-correlation surface $R(i,j)$ is defined as

$$R(\hat{i},\hat{j})\geq R(i,j)=\sum_{m=1}^{M}\sum_{l=1}^{M} f(l,m)g(l+i,m+j) \tag{29}$$

for $1<i,j<L-M+1$

where g is the search area of size LxL, f is the MxM window template representing a small image block and $(\hat{i},\hat{j})$ is the location of maximum R(i,j).

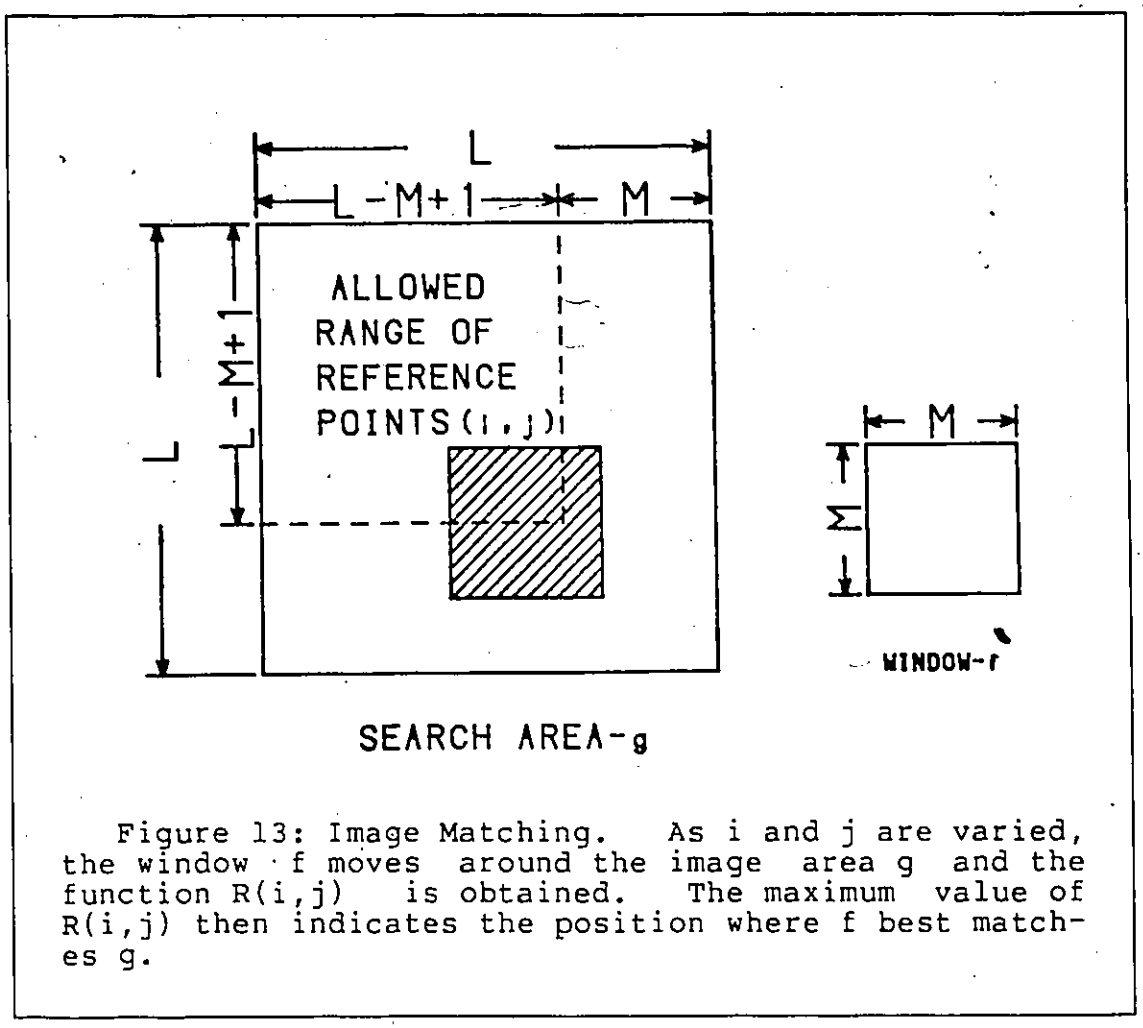


Figure 13: Image Matching. As i and j are varied, the window f moves around the image area g and the function R(i,j) is obtained. The maximum value of R(i,j) then indicates the position where f best matches g.

The motion estimation procedure, therefore, involves obtaining the cross-correlation surface R(i,j) and searching for the peak value. However, the stationary assumption of correlation is violated [3] even for a simple case where

$$g(l+i,m+j)=\max_{l,m} f(l,m)=f_m=\text{constant}$$

$$\text{for all } (l,m) \qquad (30)$$

With the actual location of match being located at shift (i',j'), it is clear that

$$R(\hat{i},\hat{j})=f_m \sum_{l=1}^{M}\sum_{m=1}^{M} f(l,m)\geq R(i',j') \qquad (31)$$

Therefore, the search does not necessarily yield the correct registration point. In order to improve the performance of the search, the correlation surface is normalized as

$$R_n(i,j) = \frac{\sum_{l=1}^{M}\sum_{m=1}^{M} f(l,m)g(l+i,j+m)}{\{[\sum_{l=1}^{M}\sum_{m=1}^{M} f^2(l,m)][\sum_{l=1}^{M}\sum_{m=1}^{M} g^2(l+i,m+j)]\}^{\frac{1}{2}}} \leq 1 \qquad (32)$$

$$\text{for } 1<i,j<L-M+1$$

By the Cauchy-Schwartz inequality, this takes on the maximum value when $f(l,m)=g(l+i,m+j)$. The normalization factor in the denominator, therefore, has a tendency to sharpen the true peak of the cross-correlation surface.

Besides the correlation function, other similarity measures exist, and two commonly used measures can be found in Table 3. A comparison [64] shows that the use of correlation coefficient and sum of absolute difference both yield

results superior to those using the unnormalized correlation [3].

TABLE 3

Similarity Measure Expressions

Correlation Coefficient $$P_{i,j} = \frac{\sum_{l=1}^{M}\sum_{m=1}^{M}[f(l,m)-\bar{f}][g(l+i,j+m)-\bar{g}]}{\{[\sum_{l=1}^{M}\sum_{m=1}^{M}\{f(l,m)-\bar{f}\}^2][\sum_{l=1}^{M}\sum_{m=1}^{M}\{g(l+i,m+j)-\bar{g}\}]\}^{1/2}}$$
Sum of Absolute Difference $$a_{i,j} = \sum_{l=1}^{M}\sum_{m=1}^{M}\lvert f(l,m)-g(l+i,j+m)\rvert$$

Furthermore, if some image preprocessing is performed prior to the search, additional improvement can be obtained. The preprocessing step tends to tailor the images to better suit the similarity model. Among all preprocessing schemes, it has been found [64] that the highest performance is obtained when the search is done on the image gradient instead of the original image.

An alternate approach to the computation of the image cross-correlation function is through the use of the Fourier transform [3]. The well known convolution theorm of Fourier

analysis states that convolution in the time or space domain is equivalent to multiplication in the frequency domain. Since correlation is a form of convolution, an alternate method of computing the correlation function exists and can be implemented using the fast Fourier transform involving much less computation.

3.2.3 The Method of Differentials

The basis of this method is to relate the spatial and temporal information of an image sequence. In mathematical terms, it is nothing more than a truncated version of the Taylor series. If the intensity function of an image sequence is assumed to be continuous, its Taylor series expansion about a point (x,y,t), taking only the linear terms, will be given by

$$I(x+dx,y+dy,t+dt)=I(x,y,t)+I_x dx+I_y dy+I_t dt \qquad (33)$$

The higher order terms have been ignored because dx, dy and dt are small.

If after a time duration dt, a translation (dx,dy) of the image is observed, the above equation will yield

$$-I_t = I_x \frac{dx}{dt} + I_y \frac{dy}{dt} \qquad (34)$$

In a discrete real-time image sequence, I_t can be approximated by the temporal difference. Similarly, I_x and I_y can be approximated by the horizontal and the vertical image gradient respectively. Assuming that the images are acquired at unit time intervals, the above equation reduces to

$$\Delta I(x,y) = -\Delta x I_x(x,y) - \Delta y I_y(x,y) \tag{35}$$

Thus, by calculating ΔI and (I_x, I_y) at two spatially nearby points, a unique solution to $(\Delta x, \Delta y)$ will be obtained.

However, image noise will cause estimation of the gradient function to be in error. Let the error in the gradient estimation I_x, I_y and I_t be e_x, e_y and e_t. Then, for a true displacement of $v=(X_q, Y_q)$, the minimum distance (see Figure 14) from v to the estimated constraint line is

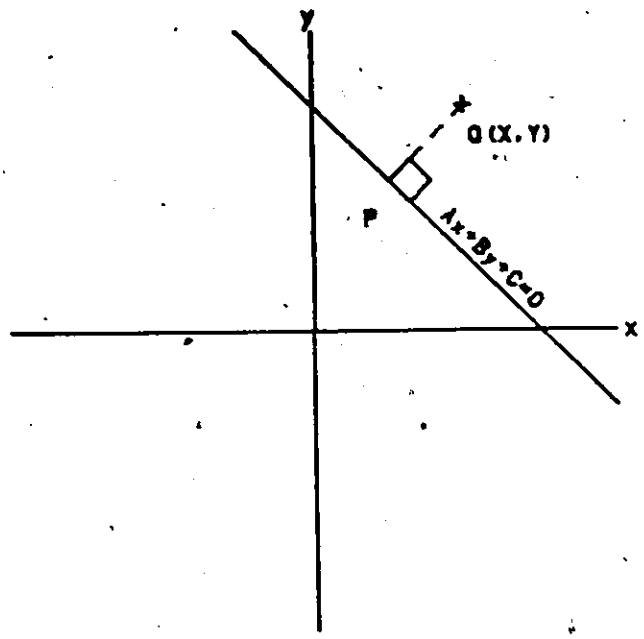
$$e = \frac{(I_x+e_x)X_q + (I_y+e_y)Y_q + (I_t+e_t)}{[(I_x+e_x)^2 + (I_y+e_y)^2]^{1/2}} \tag{36}$$

Since it is known that $xI_x + yI_y + I_t = 0$, the minimum distance is therefore

$$e = \frac{X_q e_x + Y_q e_x + e_t}{[(I_x+e_x)^2 + (I_y+e_y)^2]^{1/2}} \tag{37}$$

From the equation, it is noticed that

1. the perturbation increases with the magnitude of the displacement vector $\underline{v}$. In other words, the accuracy of the estimation decreases as the actual displacement or shift increases.



$$\overline{PQ} = \frac{AX + BY + C}{(A^2 + B^2)^{1/2}}$$

Figure 14: The effect of gradient estimation error on the method of differentials. In a noiseless situation, the constraimt line should pass through the point Q. However, with estimation error in the gradient terms, the line has a minimum distance from Q given by PQ.

2. the perturbation decreases as the magnitude of the spatial gradients increase but is unaffected by the magnitude of the temporal gradient I_t, and
3. the magnitude of the perturbation is bounded by the ratio of the error in the temporal difference to the magnitude of the spatial gradient.

The solution obtained using only two neighbourhood points is therefore not reliable. In practice, the solution to equation 35 is obtained through linear regression using more than two neighbouring points. For each pixel, a point (I_x, I_y, I_t) in the 3 dimensional space is defined. The least square estimate to equation 35 is therefore obtained by finding the equation of the geometric plane which best fits the data points. In Appendix B, it is shown that the least square estimate can be derived as

$$\mu = \frac{(\Sigma I_y^2)(\Sigma I_x I_t) - (\Sigma I_x I_y)(\Sigma I_y I_t)}{(\Sigma I_x I_y)^2 - (\Sigma I_x^2)(\Sigma I_y^2)}$$

$$\nu = \frac{(\Sigma I_x^2)(\Sigma I_y I_t) - (\Sigma I_x I_y)(\Sigma I_x I_t)}{(\Sigma I_x I_y)^2 - (\Sigma I_x^2)(\Sigma I_y^2)} \qquad (38)$$

where μ and ν are the horizontal, and vertical displacement estimates, respectively.

It should be noted that the linear relationship of equation 35 is not always satisfied in real images. At sharp transitional edge regions of the image where the grey levels change rapidly (ie. very large gradients) and the transition zone is narrow, a shift larger than the width of the transition zone will result in a violation of equation 35 (see Fig 15). A solution to the problem is by pre-smoothing the images before the calculation of the gradient and the frame difference. In this pre-smoothing step, each pixel is replaced by the local neighbourhood average ,thus reducing the

error terms e_x, e_y, e_t and preventing narrow transition zones. However, because of the amount of computation required to smooth an image digitally, the process is often slow.

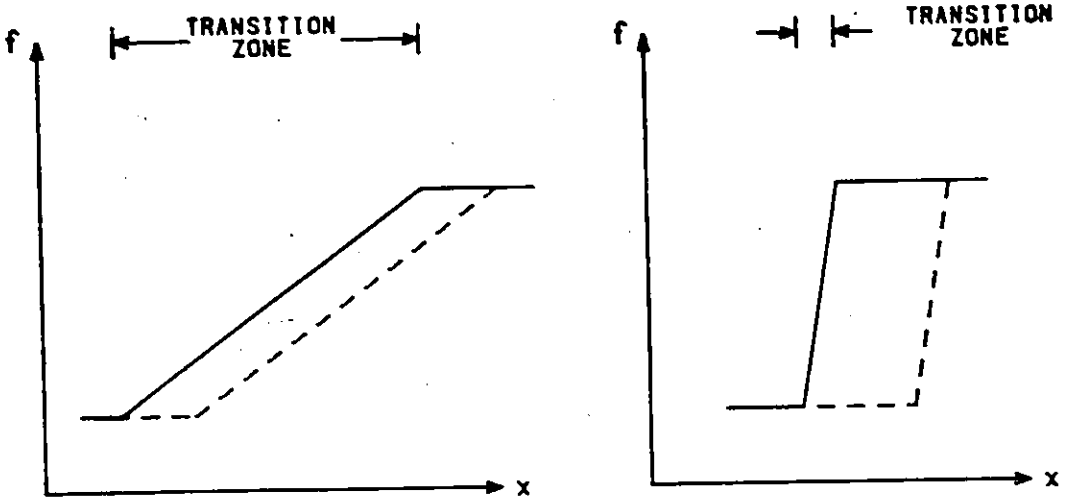


Figure 15: Effect of edge sharpness on motion estimation. a) Sharp edge: with the shift greater than the width of the transition zone, the amount of displacement cannot be measured using the gradient function. b) Smooth edge: with the shift much less than the width of the transtion zone, there is a direct relationship between the difference and the gradient values.

An alternative to image smoothing is to perform the displacement estimation upon the previous frame estimate. With the fact that the perturbation decreases as the magnitude of the displacement vector $\underline{v}$ decreases, a recursive approach

can be adopted to refine the estimation through finite iterations. For example, if the true shift between the (k-1)-th and the k-th frames is 4 pixels and the first estimation is 3 pixels, then by shifting back the k-th frame 3 pixels, a second interation may be, say, 0.9 pixel. The final estimation is then equals to 3+0.9 = 3.9 pixels. This iterative process converges to the true value but eventually the error in the temporal difference (such as the presence of quantum and video noise in the difference image of DSA) dominates the expression 35 and no further improvement is expected.

Limb and Murphy [42,41] have proposed an algorithm for estimating the speed of a moving objects in television signals. (This is similar to the case of image displacement with the exception that both moving and stationary areas are always present in a single frame.) In the algorithm, the scene is first segmented into stationary and moving areas. Within the moving area, the horizontal component of the velocity is estimated by dividing the sum of the absolute temporal differences by the sum of the absolute horizontal gradients. An intuitive illustration of the idea is shown in Figure 16 where the area and the heights of the shaded regions are correspondingly measured by the sum of the temporal differences and the sum of the horizontal gradients. By dividing these two values, a displacement estimation is obtained as shown in equation 39.

$$|d| = \frac{\Sigma|I_t|}{\Sigma|I_x|} \tag{39}$$

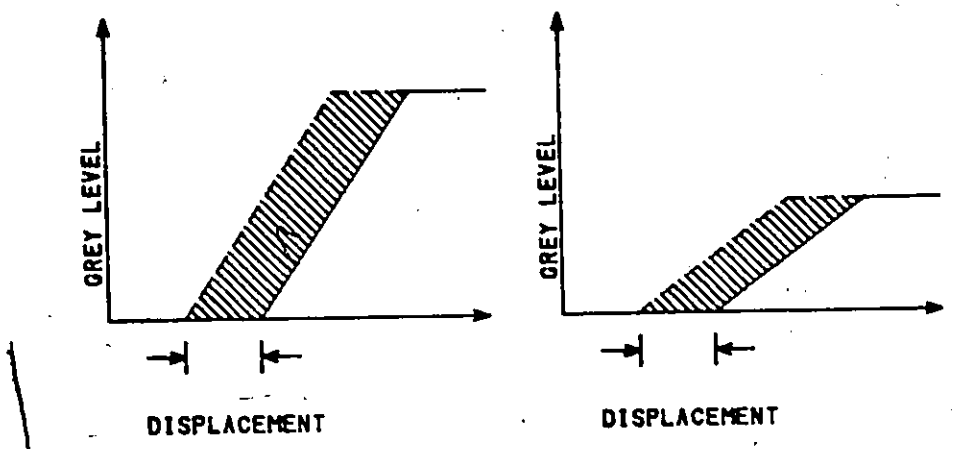


Figure 16: Illustration to the approach adopted by Limb and Murphy. The dashed line represents the one-dimensional luminance function of a moving object, the solid line shows the function in the succeeding frame. The frame difference are summed over the shaded area and represents approximately a measure of the area of a parallelogram. Dividing this value by the height of the parallelogram (ie. the sum over the horizontal element differences), a displacement estimation is obtained. The height and area of the parallelogram in b) are doulbe that of a) and, therefore, yield the same estimate.

Cafforro and Rocca [9] have proposed another displacement estimate algorithm for small image segments by simplifying the least square estimate solution. Assuming that the approximation

$$\frac{\Sigma I_x I_t}{\Sigma (I_x)^2} = \frac{\Sigma I_t \ \mathrm{sign}(I_x)}{\Sigma |I_x|} \tag{40}$$

is valid and neglecting $\Sigma I_x I_y$ terms, Cafforro and Rocca further simplify equation 38 to

$$\mu = \frac{\Sigma I_t \ \mathrm{sign}(I_x)}{\Sigma |I_x|} \tag{41}$$

so that the program or hardware realization can be simplified. The term $\Sigma I_x I_y$ is neglected is because many edges in real-world images are oriented along either the vertical or the horizontal axis, thus I_x and I_y are uncorrelated.

Both Limb and Cafforro have reported satisfactory result up to a displacement of 2.5 pixels per frame. Above this rate, the accuracy of the estimation decreases tremendously. Image pre-smoothing is required in order to improve the estimation for larger displacements.

Netravali and Robbins [51] studied the recursive estimation approach and obtained the following recursive formula:

$$D_i = D_{i-1} - e \, DFD_{m,n} \cdot \nabla I_{k-1}(m-\mu_{i-1}, n-\nu_{i-1}) \tag{42}$$

where D_i and D_{i-1} are the displacement estimates at the i-th and (i-1)-th iteration respectively, $DFD_{m,n}$ is the displaced frame difference obtained by subtracting from frame k the shifted version of frame k-1, ∇I is the spatial gradient, and e is a convergence control paremeter.

The method of differentials can also be studied from a vector point of view. In terms of vector space, ΔI can be considered as a dot product of the displacement vector $\underline{v}$ and the gradient vector $\underline{G}$

$$\begin{aligned} \Delta I &= -\underline{v}\cdot\underline{G} \\ &= -|v||G|\cos(\theta_G - \theta_v) \end{aligned} \tag{43}$$

where

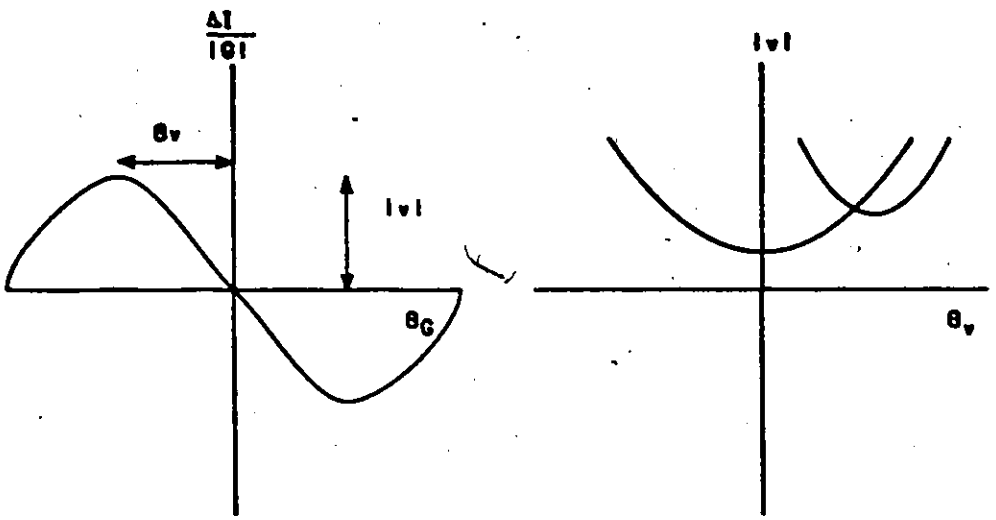
$\underline{v}=(\Delta x,\Delta y)$

$\underline{G}=(I_x,I_y)$

θ_v, θ_G are the phase of $\underline{v}$ and $\underline{G}$ respectively

For each image point, ΔI and $\underline{G}$ can be used to trace a curve through the $(|v|,\theta_v)$ space corresponding to possible displacement vectors. If the direction and magnitude of the gradient varies at different points of the image, the intersection of these curves will then uniquely specify $\underline{v}$ (see Fig 17).

Fennema and Thompson [16] have proposed an algorithm, in the vector space approach, for determining velocities of moving objects in image sequences. Instead of solving the intersection problem analytically, a Hough transform approach is used where the traced curve is digitized onto an accumulator grid in which each accumulator corresponding to a point on the curve is incremented. When the curves corre-



$$\frac{\Delta I}{|G|} = -|v|\cos(\theta_G - \theta_v) \qquad |v| = -\frac{\Delta I}{|G|\cos(\theta_G - \theta_v)}$$

Figure 17: Geometrical solution to the method of differentials. a) solution obtained by plotting data points on the $(\Delta I/|G|, \theta_G)$ space. b) solution obtained by tracing different curves on the $(|v|, \theta_v)$ space and finding the intersection.

sponding to all the points in the picture have been entered, a peak in the accumulator array corresponds to a velocity. In the process, the images are first low pass filtered and only points for which the frame difference is sufficiently large and the gradients at frame k and k+1 satisfy the restrictions

$$\{|\underline{G}(k+1)|-|\underline{G}(k)|\} < a1$$

$$|\theta_G(k+1)-\theta_G(k)| < a2 \qquad (44)$$

a1,a2 are some constants

are considered. Their results have shown that accuracy within about 1 pixel/frame can be obtained for a velocity up to 4 pixels/frame.

3.3 SUMMARY

The ability to correct motion artifact in real time can greatly improve the motion limitation of Digital Subtractive Angiography. For this purpose, a real time automated remasking scheme is proposed here consisting of a real-time motion detector which updates the mask frame whenever significant motion is detected. In addition, for the proper functioning of the real time automated remasking scheme some criteria must be set to distinguish significant motion from other temporal changes such as vascular opacification.

The problem of motion detection has been studied from a signal processing point of view and the three most commonly used techniques are: the Fourier Method, Image Matching and the method of differentials. The Fourier method makes use of the phase information of the Fourier transform. However, coherent analysis of the technique is difficult in all but the most straightforward examples. In image matching, the

position of maximum similarity is searched for between an image template and another temporally different image. Because of the amount of computation required to complete the search, the process is slow. In many applications, tradeoffs between accurarcy and complexity exist. The method of differentials is of particular interest in our DSA application because in most cases, image transforms or matching procedures are too complex for real-time hardware implementation. This algorithm relates the spatial and temporal information of an image sequence and gives best results when the objects have smooth edges and contain no prominent texture (to minimize error in the gradient terms). This gradient based method may be applied for the displacement estimate of individual pixels, small image segments, or the entire moving area and may be recursive or non-recursive. The error analysis presented in this chapter explains how the performance of the method of differentials can be improved through weighting (eg. thresholding) the spatial gradients and image presmoothing.

Chapter IV

A NEW APPROACH: OVERALL CONTRAST OF THE DIFFERENCE IMAGE (OCD)

In order to separate patient motion artifacts from vascular opacification in DSA images, some a priori knowledge is required. In chapter 2, it was shown how energy subtraction makes use of the unique energy dependence property of the x-ray attenuation characteristics of iodine. In the last chapter, a priori knowledge about the image (such as the magnitude of the gradients) is used to estimate the amount of image displacement. In this chapter, image statistics related to the difference operator and vascular opacification are studied in order to develop a pratical motion detection approach to be applied in the automated remasking process in DSA.

4.1 IMAGE STATISTICS

4.1.1 First Order Statistics

In general, first order statistics computed from an image provides information about its grey level population. It is basically a measure of how often each grey level occurs in an image, and the histogram provides an estimate of the grey level probability density function of the image. A unique-

ness theorem states that if the histogram p(z) of an image f is piecewise continuous and has non-zero value only in a finite number of z, then moments of all orders exist and the moment sequence (m_n) is uniquely determined by p(z). Conversely, (m_n) uniquely determines p(z). The n-th moment and its corresponding central moment are defined respectively as

$$m_n = \sum z^n p(n)dz$$

$$\phi_n = \sum (z-\bar{z})^n p(n)dz \qquad (45)$$

where $z=m_1/m_0$

Figure 18 shows that the grey scale distribution of a typical x-ray image does not generally approximate a Gaussian process. Moreover, certain textural properties and statistical measures can be obtained from the histogram function using the different order moments. The following are a few statistical measures of an image.

1. Measure of Overall Brightness

   The mean grey level of f, $z=(1/N)\Sigma zp(n)$, were $N=\sum p(z)$ is the number of points in f, is a measure of the overall brightness of f. It is a calculation of the average grey level that most of the pixel values occupy. Another such measure is the median grey level M_j which is defined such that half of the points of f are brighter than M_j and half are darker. The mode

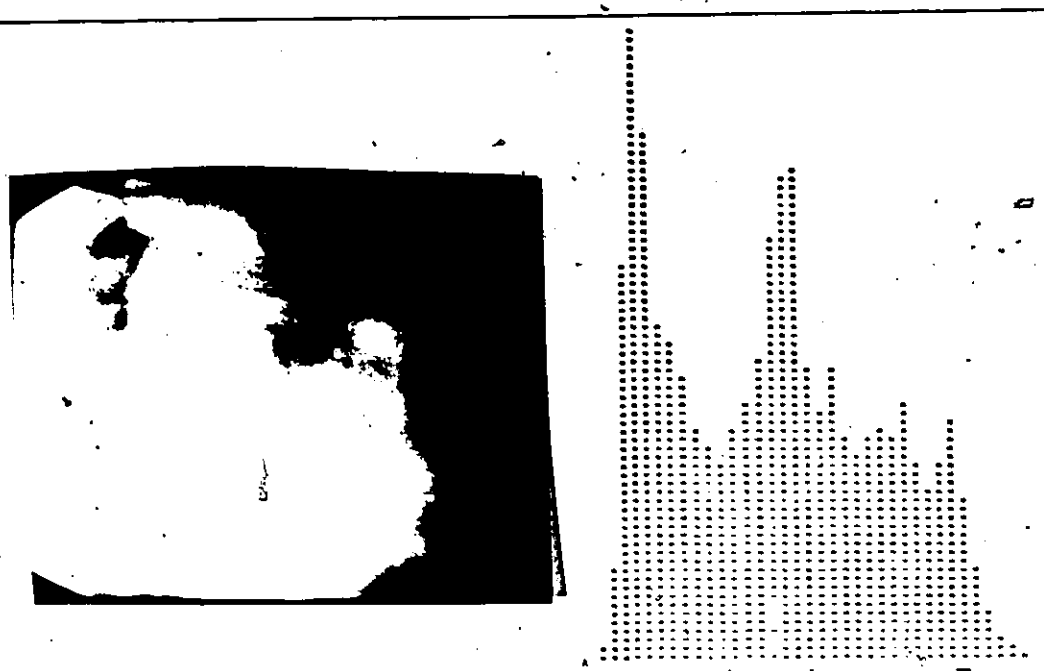


Figure 18: Histogram of a typical x-ray image. Examination of the histogram (right) from the x-ray image (left) reveals that the probability density of the image is not Gaussian. The various peaks correspond to the dark and bright regions of the image.

is a slightly different measure that indicates the grey value which occurs with the greatest frequency. The mode may not always exist, and even if it does exist, may not be unique.

2. Measure of Overall Contrast

The grey value variance of f, $\sigma^2=(1/N)\Sigma(z-\bar{z})^2p(z)$ is a measure of the overall contrast of f. While a small value indicates that the grey levels of f are all close to the mean, a large value indicates a wide range of grey levels. From the above expression,

grey values far from the mean value are given more weight in the computation. This may be of interest to the method of differentials where improved performance can be obtained through the weighting of the spatial gradient function. Another contrast measure is the interquartile, which is defined as |P75-P25| where P25 and P75 are the 25-th and the 75-th percentile for the data. Other percentiles can also be used.

3. Degree of asymmetry of the histogram distribution

Skewness is a measure of the degree of asymmetry of a distribution. If the histogram has a longer tail to the right of the central maximum than to the left, it is said to be skewed to the right or to have a positive skewness. If the reverse is true, it is said to be skewed to the left or have negative skewness. This measure is evaluated by [(mean-mode)/standard deviation]. If the distribution is unimodal (having only one mode), the value is a measure of the ratio of the occurance of bright and dark points. Another measure of skewness using the third central moment is given by

$$\frac{\phi_3}{\sigma^3} = a_3 \qquad (46)$$

4. Degree of peakness of the histogram distribution

The degree of peakness, termed kertosis, of a distribution measures the distribution of grey level around the peak. One measure of kertosis using the fourth central moment is given by

$$a_4 = \frac{\phi_4}{\sigma^4} - 3 \qquad (47)$$

which is taken relative to the normal distribution. For a normal distribution, which is not very peaked nor very flat-topped (ie. mesokurtic) $a_4=0$. Therefore, a distribution having relatively high peak (ie. leptokurtic) has $a_4>0$ whereas, the curve which is flat-topped (platykurtic) has a measure $a_4<0$.

4.1.2 Second Order grey level statistics

Although some textural information can be described using first order grey level statistics, nothing is known about the grey level arrangement of the image. The histogram p(z) remains the same no matter how the pixels are arranged. An image of half black and half white, or of checkboard, or of salt and pepper noise may have the same histogram p(z). More spatial information can be obtained by studying, for example, how often the possible pairs of grey level occurs in a given relative position.

The second order statistics are studied by forming a kxk matrix M_δ whose (i,j) element is a measure of the frequencies of occurance that a point having grey level Z_i occurs in position $\delta=(\Delta x,\Delta y)$ relative to another point having grey level Z_j , $1<i,j<k$. Elements near the diagonal of M_δ therefore correspond to pairs of grey levels that are nearly equal, while elements far from the diagonal correspond to pairs that are quite different. Studies [50,64] of the second order statistics show that the neighbouring pixels are highly correlated (equivalently, video energy of an image is concentrated in the low frequencies). As a result, the histogram of the adjacent element difference signal (eg. $[f(x,y)-f(x-1,y)]$ for $\delta=(1,0)$) is highly peaked at zero.

4.2 STATISTICAL ANALYSIS OF TEMPORAL DIFFERNCE

Because temporal changes are enhanced using image subtraction, the difference image is of particular importance in studying the temporal changes of an image sequence. In this section, statistics of the image obtained by the subtraction of two images are studied under the assumption that patient motion is essentially translational. As in other motion estimation techniques, this is a weak assumption. However, we shall adopt the model so that an analytical study is possible.

4.2.1 Difference of two images

The difference $z_3(x,y)$ that results from a temporal subtraction can be defined as

$$\begin{aligned} z_3(x,y) &= z_2(x,y) - z_1(x,y) \\ &= [I_2(x,y)+n_2(x,y)]-[I_1(x,y)+n_1(x,y)] \end{aligned} \tag{48}$$

where I_2, I_1 and n_2, n_1 are the information and the noise portion of the images z_2, and z_1, respectively. If n_1 and n_2 are uncorrelated white noise, the mean (brightness) and variance (contrast) of $z_3(x,y)$ can be obtained as

$$E[z_3(x,y)] = E[I_2(x,y)] - E[I_1(x,y)]$$

and

$$\begin{aligned} var[z_3(x,y)] = var[I_2(x,y)]+var[I_1(x,y)] \\ -2cov[I_1(x,y),I_2(x,y)] \\ +var[n_1(x,y)]+var[n_2(x,y)] \end{aligned} \tag{49}$$

With respect to DSA images, the assumption is, however, ill-conditioned. According to the Poisson nature of quantum statistics, the uncertainty in the noise variance related to x-ray detection is signal dependent (Refer to Appendix A for the derivation of the output noise in DSA) and modifications to equation 49 would have to be made.

In order to avoid the the handling of signal dependent noise, let us combine the information and the noise portion of the image together and rewrite equation 49 as

$$R_{33} = R_{11} + R_{22} - 2R_{12} \tag{50}$$

where R_{ii} is the variance of z_i for i=1,2,3

and R_{ij} is the covariance between z_i and z_j

with i,j=1,2,3

With the simplified expression, the statistical relationship between z_1, z_2 and z_3 is made explicit. It can be depicted geometrically by plotting z_2 against z_1 for each pixel. Because z_2 and z_1 are highly correlated, most samples orient along the line $z_2=z_1$ whereas the distribution of $z_3(x,y)$ can be obtained as a projection of data along the line $z_2=z_1$ as shown in Figure 19 . In DSA, if z_1 (mask) and z_2 (live) are registered correctly, the difference should have essentially zero mean with variance contributed only from the noise portion of the two images. With the contrast agent present, z_1 and z_2 will be highly correlated except at locations where the dye appears. However, if the two images are not aligned properly, the correlation between z_2 and z_1 will decrease, resulting in an increase in the variance of $z_3(x,y)$.

In addition, the covariance function between the mask and the difference can be expressed as

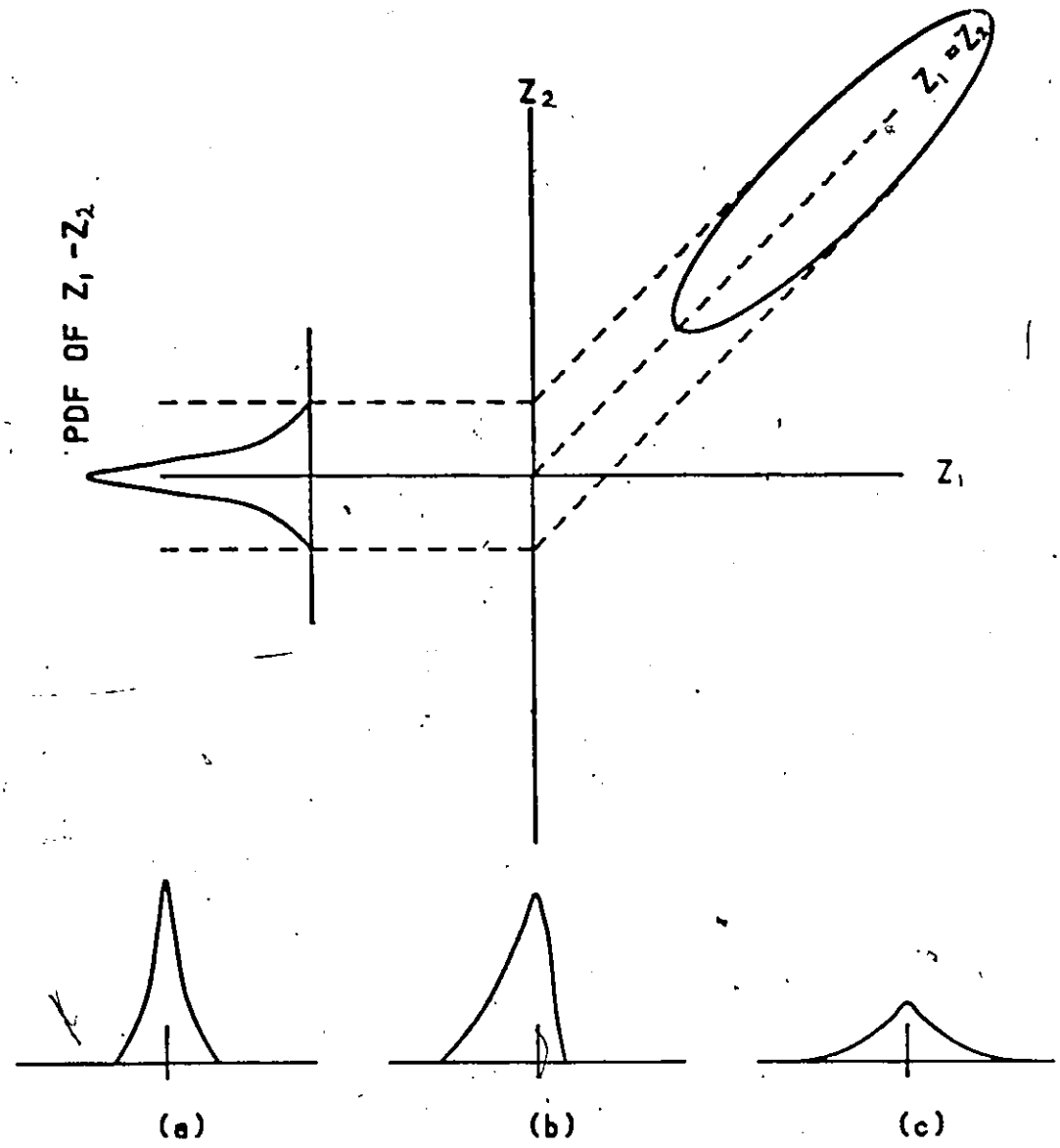


Figure 19: Image statistics involved in subtraction. Data points (z_1, z_2) projected along the line $z_1 = z_2$ contribute to the formation of the histogram of the difference image. a), b) and c) correspond to typical difference histograms resulting from noise, dye and motion respectively.

$$\mathrm{cov}\{z_3(x,y), z_1(x,y)\} = E\{[z_1(x,y) - \bar{z}_1][z_3(x,y) - \bar{z}_3]\}$$

$$
\begin{aligned}
&= \mathrm{cov}\{z_1, z_2\} - \mathrm{var}\{z_1\} \\
&= [(1/2)(\mathrm{var}\{z_1\} + \mathrm{var}\{z_2\} - \mathrm{var}\{z_3\})] \\
&\quad - \mathrm{var}\{z_1\} \\
&= (1/2)(\mathrm{var}\{z_2\} - \mathrm{var}\{z_1\} - \mathrm{var}\{z_3\}) \\
&= (1/2)[(R_{22} - R_{11}) - R_{33}] \qquad (51)
\end{aligned}
$$

In a similar fashion, the other covariance functions can be obtained and the simplified results are as follows:

$$
\begin{aligned}
R_{12} &= (1/2)[(R_{11} + R_{22}) - R_{33}] \\
R_{13} &= R_{12} - R_{11} = (1/2)[(R_{22} - R_{11}) - R_{33}] \\
R_{23} &= R_{22} - R_{12} = (1/2)[(R_{22} - R_{11}) + R_{33}] \qquad (52)
\end{aligned}
$$

One can see that, with the above equations, the covariance between z_1 and z_2 can be measured using the variance of the difference image (ie. R_{33}). The same is true for the cross-correlation functions if R represents the cross-correlation function instead of the covariance function.

4.2.2 Translational difference

Now if one further assumes that the image change between z_2 and z_1 can be modelled by translational motion $(\Delta x, \Delta y)$, then z_1 will be related to z by the relation:

$$z_1(x,y) = z_2(x+\Delta x, y+\Delta y) \qquad (53)$$

Of course one may question that with complex patient motion and with quantum and video noise inherent in both z_1 and z_2,

the translational relationship no longer holds. However, for an analytical study, we shall assume the ideal case where the translational model for motion is valid.

It can then be shown that the covariance function R_{12} between z_2 and z_1 can be computed using the autocovariance function $R_{11}(\Delta x, \Delta y)$ so that

$$\begin{aligned} R_{12} &= E\{(z_1(x,y)-\bar{z}_1)(z_2(x,y)-\bar{z}_2)\} \\ &= E\{z_1(x,y)z_2(x,y)\}-\bar{z}_1\bar{z}_2 \\ &= E\{z_1(x,y)z_1(x+\Delta x, y+\Delta y)\}-\bar{z}_1\bar{z}_1 \\ &= R_{11}(-\Delta x, -\Delta y) \\ &= R_{11}(\Delta x, \Delta y) \end{aligned} \tag{54}$$

Combining equation 54 and equation 52 yields

$$\begin{aligned} R_{12} = R_{11}(\Delta x, \Delta y) &= (1/2)[(R_{11}+R_{22})-R_{33}] \\ &= R_{11}-(1/2)R_{33} \end{aligned} \tag{55}$$

It is clear, from equation 55, that if the spatial statistics of image z_1 (the mask) are available, the amount of translational shift between z_2 and z_1 can be computed directly using the variance of the difference image. This is of particular benefit in the DSA problem. Since the difference image is readily accessible in DSA and statistical parameters of an image can be directly obtained from its histogram, the hardware implementation is greatly simplified.

4.3 SPATIAL STATISTICS

Should the image autocovariance function of z_1 be known for all points of the plane, to get a measure of relative displacement would be an easy task through the use of equation 55 . However, the measuring of such a function is a very long procedure. A feasible alternative to the estimation of the function through image modelling is outlined in this section.

4.3.1 Exponential model

The spatial autocovariance function of an image depends highly upon the detail in the picture. In general, it is maximum at zero pixel shift and decreases with increasing shift. The rate of decrease is large for shifts close to zero, but becomes smaller for larger shifts. Various reports [50,64] have suggested that the autocovariance function (and similarly the autocorrelation function) can be satisfactorily approximated by an expoential function of the form

$$R(\Delta x,\Delta y)=A \exp\{-\alpha|\Delta x|-\beta|\Delta y|\} \qquad (56)$$

where Δx and Δy are the spatial displacement and α, β and A are positive constants determined by the image texture. In general, an image of coarse texture will exhibit a higher neighbourhood correlation (hence a smaller α, β values) than

an image of fine texture. The spread of the autocovariance function is therefore larger for an image of higher textural coarseness.

4.3.2 Quadratic Function Modelling

If $\Delta x, \Delta y$ are small, one can approximate the autocovariance function $R(\Delta x, \Delta y)$ by the Taylor series expansion about the origin such that by neglecting terms higher than the second order, the following expression is obtained:

$$R(\Delta x,\Delta y)=R(0,0)+\Delta x\frac{d}{dx}R(0,0)+\Delta y\frac{d}{dy}R(0,0)$$
$$+(1/2)[\Delta x^2\frac{d^2}{dx^2}R(0,0)+2\Delta x\Delta y\frac{d^2}{dxdy}R(0,0)+\Delta y^2\frac{d^2}{dy^2}R(0,0)] \quad (57)$$

The first order terms naturally become zero, because of the peak of $R(\Delta x,\Delta y)$ at $\Delta x=\Delta y=0$. This can be shown mathematically by defining the (i,j)-th order moment of the spectral density $S(f_x, f_y)$ as

$$m(i,j)=\int\int_{-\infty}^{\infty} S(f_x, f_y) f_x^i f_y^j df_x df_y \quad (58)$$

Since the power spectral density S of a real process is even, m(i,j) is zero for (i+j) odd. Thus the first derivative terms in equation 57 is zero. When (i+j) is even

$$\frac{d^{i+j}}{dx^i\,dy^j}R(0,0)=(-1)^{(i+j)/2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}S(f_x,f_y)f_x^{\,i}f_y^{\,j}\,df_x\,df_y$$

$$=(-1)^{(i+j)/2}\,m(i,j) \qquad (59)$$

because R and S are Fourier transform pair.

The second order terms can be obtained by the equation

$$(-1)^j\frac{d^{i+j}}{dx^i\,dy^j}R(0,0)=E\left\{\frac{d^i z(0,0)}{dx^i}\,\frac{d^j z(0,0)}{dy^j}\right\} \qquad (60)$$

therefore

$$\frac{d^2}{dx^2}R(0,0)=-E\left\{\frac{dI}{dx}\,\frac{dI}{dx}\right\}$$

$$\frac{d^2}{dy^2}R(0,0)=-E\left\{\frac{dI}{dy}\,\frac{dI}{dy}\right\}$$

$$\frac{d^2}{dxdy}R(0,0)=-E\left\{\frac{dI}{dx}\,\frac{dI}{dy}\right\}$$

Equation 57 now reduces to

$$R(\Delta x,\Delta y)=R(0,0)-(1/2)\left\{\Delta x^2E\left\{\left(\frac{dI}{dx}\right)^2\right\}+\Delta y^2E\left\{\left(\frac{dI}{dy}\right)^2\right\}\right.$$

$$\left.+2\Delta x\Delta yE\left\{\frac{dI}{dx}\,\frac{dI}{dy}\right\}\right\} \qquad (62)$$

If all the covariance terms on the right hand side of equation 62 are defined, an approximation of the image autocovariance function is obtained for the main lobe where $\Delta x, \Delta y$ are small. Outside of this range, the remaining higher order terms in the Taylor series contribute to a large degree and the corresponding portion of the curve described by equation 62 deviates from the actual covariance function. Poor description is therefore expected for large values of $\Delta x, \Delta y$ (Figure 20).

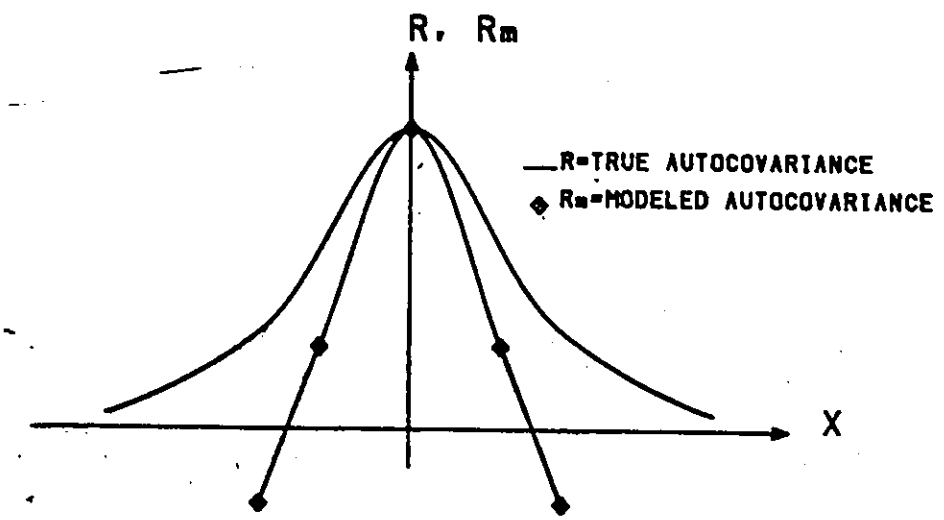


Figure 20: One dimensional estimate of the image autocovariance function. The quadratic model yields close description for the main lobe. Outside of this range, poor description is observed.

Relating to our DSA case, the quadratic function suggests that the autocovariance function of the mask, z_1, can be ob-

tained simply with measures of the variance of its gradients. Since the gradient operator can be implemented in real time using only a delay circuit and a subtraction operation, with the model the autocovariance function of the mask image can be obtained as soon as the mask is acquired.

4.4 MOTION ESTIMATION

In the previous section, it is shown that the image autocovariance function can be described by the quadratic expansion model whose parameters are obtained by measuring the variance of the horizontal and the vertical luminance gradients. The same model, but for a spatial translation, can also be assumed for the covariance function between two temporally different images z_1 and z_2 and is shown explicitly in equation 54.

If equation 62 is combined with equation 55, the variance of the difference image resulting from translational motion can be related to the spatial statistics of z_1 by

$$\begin{aligned} R_{33} &= 2[R_{11} - R_{11}(\Delta x, \Delta y)] \\ &= 2[R_{11} - R_{11} \\ &\quad + (1/2)(\Delta x^2 \mathrm{var}\{I_x\} + \Delta y^2 \mathrm{var}\{I_y\} + 2\Delta x \Delta y \mathrm{cov}\{I_x, I_y\})] \\ &= \Delta x^2 \mathrm{var}(I_x) + \Delta y^2 \mathrm{var}(I_y) + 2\Delta x \Delta y \mathrm{cov}(I_x, I_y) \end{aligned} \tag{63}$$

In fact, this result can be derived from the Taylor expression of the method of differentials where

$$\Delta I(x,y) = -\{\Delta x I_x + \Delta y I_y\} \tag{64}$$

so that the mean and variance of the difference, which describes the overall brightness and contrast, can be computed as

$$E\{\Delta I(x,y)\} = -[\Delta x E\{I_x\} + \Delta y E\{I_y\}]$$

and

$$\text{var}\{\Delta I(x,y) = \Delta x^2 \text{var}\{I_x\} + \Delta y^2 \text{var}\{I_y\} + 2\Delta x \Delta y \text{cov}\{I_x, I_y\} \tag{65}$$

In order to obtain a displacement estimate $(\Delta x, \Delta y)$, one would have to solve at least two simultanous equations. In the automated remasking scheme, however, the direction of motion is not of concern and, unlike the image matching technique, no image re-registration is required. The estimation problem may therefore be simplified. If we assume that the covariance term $\text{cov}\{I_x, I_y\}$ is orders of magnitude lower than the others, and that the variance of the vertical gradient and the horizontal gradient are approximately equal, then a measure of the amount of displacement d can be obtained with

$$\text{var}\{\Delta I\} = (\Delta x^2 + \Delta y^2)\text{var}\{I_x\}$$

$$d^2 = (\Delta x^2 + \Delta y^2) = [\text{var}\{\Delta I\}]/[\text{var}\{I_x\}]$$

$$d = [\sigma\{\Delta I\}]/[\sigma\{I_x\}] \tag{66}$$

This is similar to the approach proposed by Limb and Murphy [42,41], except that the summing operator is replaced by the standard deviation. Since the temporal difference and the gradient are weighted according to their magnitude, a better estimate is expected with the OCD approach.

The algorithm for the OCD approach to motion estimation can be summarized as follows:

1. A quadratic expansion model is assumed for the image autocovariance function,
2. The parameters for the model are obtained by measuring the variance and and covariance factors of the horizontal and vertical gradients,
3. The same model, but for a spatial translation, is assumed for the covariance function between the two images.
4. The displacement estimation is calculated using equation 66.

Having derived an approach for motion estimation to be applied to DSA, there remains a signal detection problem of separating the difference effects due to motion and that due to the changes in vessel opacification. As studied in Chapter 2, the degree and orientation of motion varies from case to case. It generally produces in the difference image high contrast artifact consisting of bright and dark streaks. From a medical information point of view, the amount of mo-

tion is considered significant if the resulting artifact severly impairs the diagnostic information present. Therefore, in the examination of fine artery branches, even a slight motion is not to be tolerated. A simple example may help to illustrate the point. For example, suppose that the maximum allowable motion in the subtraction is 20% of the smallest vessel which is to be visualized. Then, for an image intensifier of 20 cm input and a matrix size of 512x512, a motion of 1 pixel is the maximum allowable if a 2mm artery is to be visualized clearly (This is computed by 512/2/200x20%). Motion of more than 1 pixel will, therefore, result in significant artifact. This measure can be relaxed a bit if the vessel of concern has a diameter larger than 2mm.

4.5 CONTRAST AGENT IN DETECTION

The principal concept developed in the OCD approach is that by assuming a translation model of motion, an estimate of the amount of displacement can be obtained by measuring the variance of the difference and the mask image. One then considers motion to be significant if the estimate exceeds a predetermined threshold. However, in the case of image change due to contrast opacification, the translational model will not be appropriate to describe the change. Statistical change due to contrast agent may cause an increase in the contrast, and thus the variance, of the difference im-

age. As a result, a large estimated displacement may be obtained according to equation 66, even though no motion is present. This is, in particular, true for a high contrast vessel opacification image. Therefore, if the detection problem is not handled properly, false detection of motion may result. In order to prevent this, more a priori knowledge of the effect of dye on images will be required.

Consider image change due to vessel opacification. Since the live image with iodinated agent usually contains lower intensity (ie. with dye in a vessel) than the corresponding mask image, the subtraction of the mask from the live frame will generally yield more negative pixels than positive pixels. The grey scale histogram of the difference image will therefore exhibit an asymmetric distribution with a longer tail on the negative side. The skewness of the histogram depends upon the concentration of contrast agent and the amount of area opacified. For example, if the histogram of the difference image has a large spread or variance due to the large value of negative numbers coupled with the subtraction of a high concentration of dye, a large skewness to the left or negative skewness will be expected. The amount of skewness decreases as the concentration of contrast agent and the area of opacification decreases. On the other hand, as the concentration of contrast agent or the area of opacification decrease, the variance of the difference also decreases and a small estimation of displacement will result.

For standard misregistration artifacts, however, movement of a high contrast object is generally associated with both positive and negative differences, and a histogram showing symmetry is expected. Therefore, assuming that a large motion estimation according to equation 66 is caused by either large motion or contrast of high concentration, an additional constraint to ensure that a large variance of the difference is caused by motion and not by contrast agent is that the histogram of the difference image exhibits a symmetrical property. There remains a question, however, concerning the level of iodine contrast and the amount of opacification area required in order that a reasonable negative skewness is obtained in the case of low contrast vessel opacification. In the next chapter, simulation results on the performance of the OCD approach are shown and different statistical measures are studied in order to minimize the possibility of false detection of motion.

4.6 SUMMARY

In this chapter, image statistics associated with the mask and the difference image are studied and a statistical approach to motion detection is presented. A model for the image autocovariance function of the mask is studied. The cross covariance function between the mask and the live images are then assumed to act according to this model. The amount of motion is computed using the variance of the dif-

ference image and is considered to be significant if the variance is greater than a threshold whose value depends on the parameters of the model. In order to prevent false detection due to contrast agent, the skewness of the difference image is computed. A final decision is made according to the skewness and the variance of the difference image and the hypothesis of significant motion is accepted if the variance exceeds the threshold and no significant negative skewness is observed. Otherwise, the hypothesis that significant motion has occured is rejected. The acceptance of the above hypothesis would result in an update of the mask image in the proposed automatic re-masking algorithm.

Chapter V

RESULTS

5.1 NATURE OF THE TESTED IMAGES

The experiments throughout this chapter were carried out on sequences of angiogram images prepared by the CGR company of France. Each image sequence is carefully sampled to accommodate the arrival and the wash out of iodine agent and each image is sampled in an array of 512 by 512 pixels, and quantized with 8 bits/pixel (corresponding to a grey level range from 0 to 255). Image sequence BROWN consists of seven frames of images obtained from an intravenous carotid angiogram examination with the patient swallowing occuring between frame 1 and frame 2. Image sequences GREY, WHITE and BLACK are examinations of the aortic arch, coronary arteries, and intracranial vessels, respectively, and were used in the study of the effect of iodine contrast on the detection algorithm. The amount of patient motion in each of these three sequences is minimal, but the level of contrast opacification is different in each case.

The effect of patient swallowing is shown in Figure 21, which was obtained from sequence BROWN by subtracting frame 1 (the mask) from frame 6 (maximum opacification) using

Figure 21: Motion Artifacts from BROWN. Due to the swallowing artifacts, carotid arteries are not perceptible

FORTRAN programs developed as part of this research and using the University of Ottawa, Department of Electrical Engineering VAX timesharing computer system. The difference signals have been offset by a constant value in order to avoid negative grey values on the display. As can be seen, due to the high contrast artifact caused by swallowing, carotid arteries which supply blood to the cerebral organs are not perceptible. Figure 22 shows the result obtained through image re-registration using the normalized correlation coefficient. In the process, a search procedure was performed prior to the subtraction to realign frame 1 with frame 6 and a maximum coefficient was found at a displace-

ment of (2,-1). In spite of the attempt, very little improvement was obtained since the change that had occured was not homogeneous. Image re-registration would only reduce artifacts in certain parts of the image and introduce, on the other hand, additional artifacts in some other areas.

Figure 22: Motion Artifacts after image re-registration. The swallowing artifacts are still present and the opacification of vessels is poor.

A better solution through the use of the remasking technique is demonstrated in Figure 23 which was obtained by subtracting frame 2 from frame 6. One can see that diagnostic quality has been significantly improved and the swallowing artifacts are removed. Detection of occulsion or narrowing of blood vessel would therefore be possible.

Figure 23: Improvement of image quality obtained through remasking. Motion artifacts have been significantly reduced and diagnostic quality have improved.

A similar result is expected if this mask updating process is done in real-time with the additional motion detection unit. The effect of various factors related to the performance of the motion detection algorithm are presented later in this chapter.

5.2 IMAGE STATISTICS

In the derivation of equation 66, it has been assumed that the covariance between the horizontal and vertical gradients are small and that the variances of the horizontal and vertical gradient are comparable in magnitude. The validity of the assumption is examined in this section.

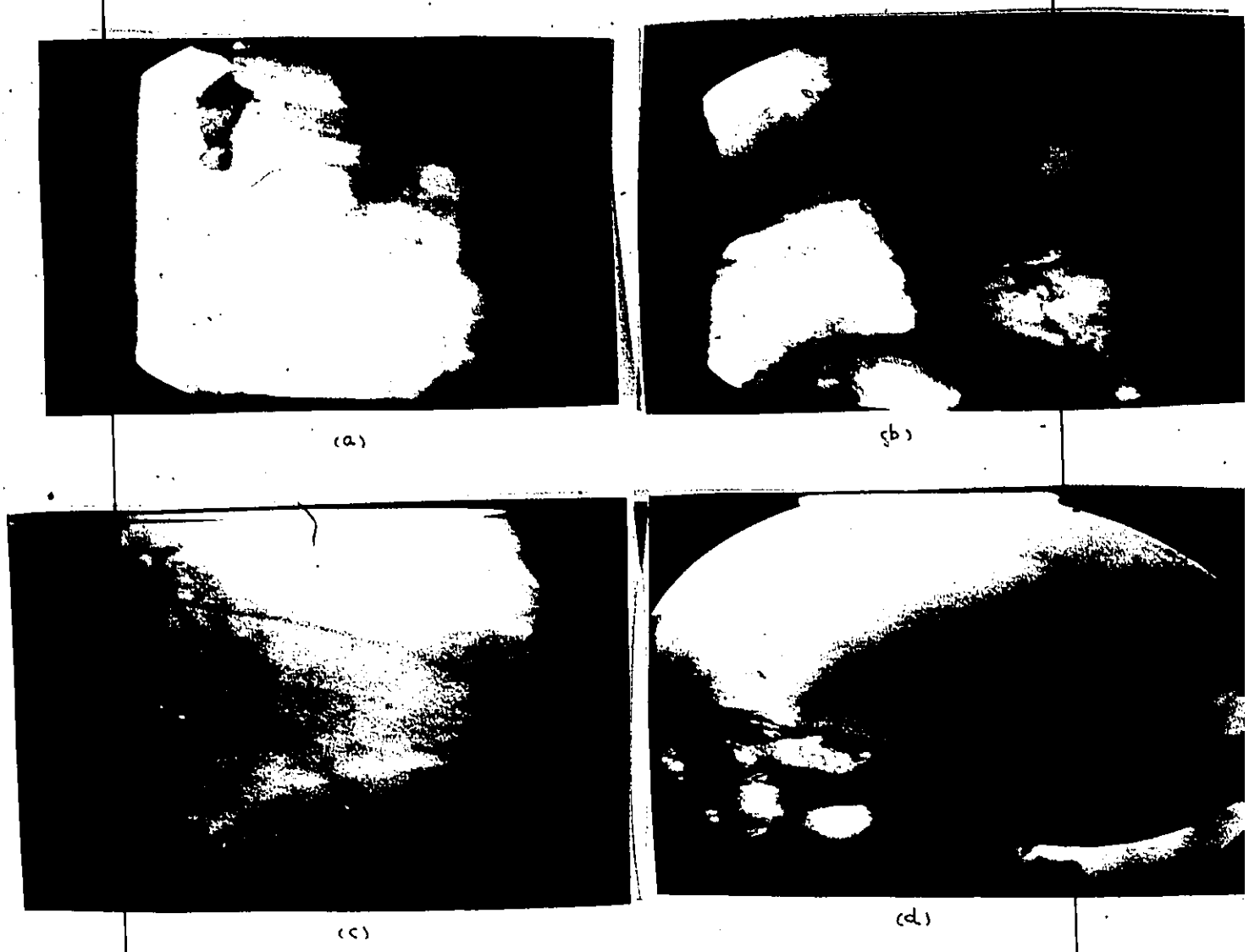


Figure 24: Intravenous angiograms before temporal subtraction. a) BROWN: the neck area before the iodine opacification of carotid arteries, b) GREY: chest region of a patient before the opacification of the aortic arch, c) WHITE: image of the left ventrical, d) BLACK: image of the skull before the opacification of intracranial vessels.

The simulations were performed on the various images shown in Figure 24. Each image represents the first frame of the angiogram sequences mentioned earlier. The horizontal and the vertical gradients are then computed respectively using the expressions:

$$I_x = \frac{I(i+1,j)-I(i-1,j)}{2} \quad \text{and} \quad I_y = \frac{I(i,j+1)-I(i,j-1)}{2} \qquad (67)$$

The centered gradient operator prevents the current pixel from being used in both the computation of the vertical and the horizontal gradient (eg., $I(i+1,j)-I(i,j)$ and $I(i,j+1)-I(i,j)$.) and thus avoids a biased gradient estimation. The variance and covariance of the gradient functions are computed for an arbitrary central region of size 256x256 and the results are shown in Table 4. The results show that the covariance terms are indeed an order of magnitude smaller than the variance terms. The variances of the vertical and the horizontal gradients, however, differ by some amount. Furthermore, it appears that the variance of the horizontal gradient is smaller than the variance of the vertical gradient for each of the tested images.

TABLE 4

Variance and Covariance of gradients

	$var(I_x)$	$var(I_y)$	$cov(I_x, I_y)$
BROWN	2.68	3.09	-0.13
GREY	3.37	4.34	-0.24
WHITE	0.48	0.58	0.06
BLACK	1.73	3.53	0.13

5.3 A COMPARISON OF DIFFERENT MOTION ESTIMATION ALGORITHMS

In order to study the performance of the OCD approach, the least square estimate (LSE) (see Appendix B) is tested as a comparison. The experiments were carried out with images generated on the computer by shifting frame 1 of BROWN[4] one pixel at a time. The temporal difference images were obtained by subtracting frame 1 (BROWN1) from its shifted versions and the centered gradient operation was used for the estimation of the spatial gradients. Image statistics were then measured for a block of 256 by 256 pixels, corresponding to the central area of the difference image and the displacement estiamtes were computed, according to the LSE, using equation 38.

---

[4] Frame 1 of BROWN is referred as BROWN1 for simplicity

Figure 25 shows the result of the estimation using the LSE method. The true displacements are the grid points of the rectangular grid and the estimates are the superimposed pattern connected with dotted lines. The result shows that for a displacement larger than 2 pixels, the algorithm breaks down due to the effect of a sharp edge and the error in the estimation of the spatial and temporal gradients.

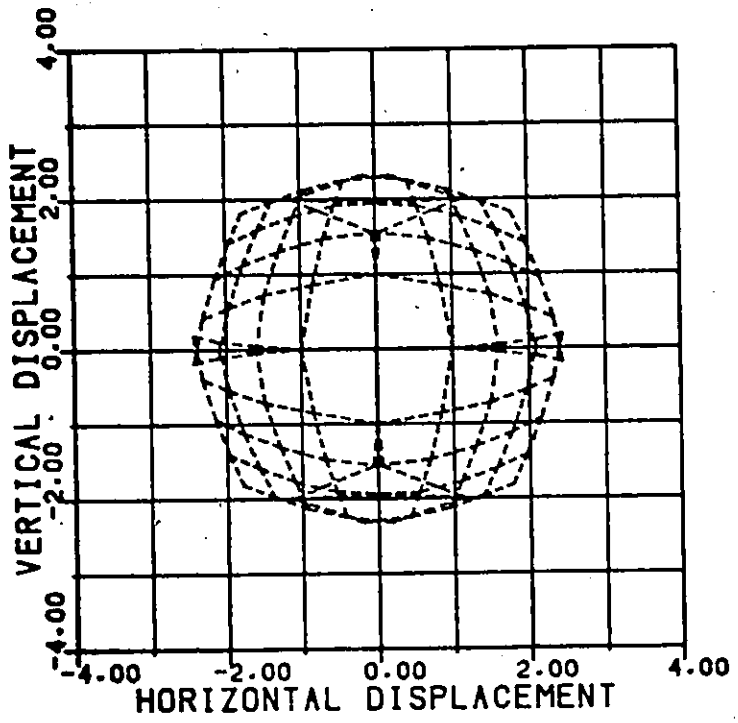


Figure 25: Performance of the LSE aprroach to motion estimation. The rectangular grid shows the true displacement and the intersections of the dotted lines are the estimated displacement. Good estimate can only be obtained for a 1 pixel displacement.

A comparison of the LSE, the Limb and Murphy's approach (LIMB) and the OCD approach is shown in Figure 26, where the displacement estimates are shown for a purely horizontal motion. In the figure, the horizontal axis represents the actual shift values and the vertical axis the estimated shift values. The result shows that both the LIMB and the OCD approach are superior to the theorectical LSE of equation 38. The poor performance of the LSE is due to the great number of image points that do not satisfy the linear relationship of equation 35 in practice. In the LIMB and the OCD approach, the estimate is obtained using the overall statistics of the images instead of a pixel based algorithm, and better estimates are obtained.

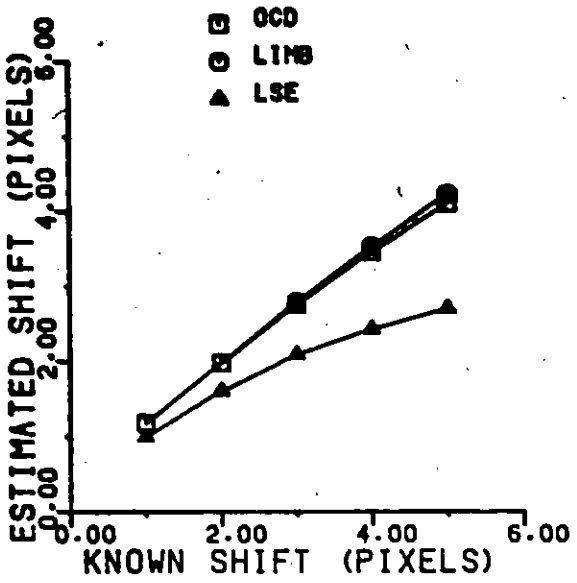


Figure 26: A comparison of the LSE, LIMB and the OCD approach in a purely horizontal displacement case. LIMB and OCD measure the overall statistics of the difference images and yield superior results to LSE.

5.4 PERFORMANCE OF THE OCD APPROACH

There exist various practical factors such as noise, edge details and textural information, which affect the performance of the OCD approach. In this section, simulations of the OCD approach to motion estimation is done to illustrate the effect of each factor. A block was taken from images generated from BROWN1 and the variances of the gradients were computed. The temporal difference images were obtained by subtracting the images from their shifted versions as mentioned above. Motion estimation was then performed by

dividing the variance of the temporal difference by the variance of the horizontal gradient.

5.4.1 Noise

The effect of image noise was studied by adding random Gaussian noise to BROWN1 before image shifting and then applying the estimation algorithm described above. It is expected that the estimation error will increase as the amount of noise increases. Image noise introduces an increase in the variance of both the difference image and the gradients. As a result, as the amount of noise increases the sensitivity of the algorithm to motion decreases. Figure 27a shows the results for different amounts of noise in the image. The block used corresponds to the central area of the image with a block size of 256 by 256 pixels. The standard deviations of the noise used were 4, 8 and 12. The increased noise was shown to cause a decrease in the accuracy of estimation. The over-estimation at 1 pixel shift is due to the use of a centered gradient operator which yields the best estimate at the 2 pixel shift with $I_x = [I(i+1,j)-I(i-1,j)]/2$. An alternative presentation of the result is given in Figure 27b where the vertical axis represents the relative estimation computed by the ratio of estimated shift (Se) over a known shift (Sk). With the 4 pixel shift case, it is illustrated how even a slight amount of noise can greatly affect the accuracy at large shifts. How-

ever, this can be improved through image pre-smoothing. As the quality of the imaging systems for DSA improves, the effect of image noise will only play an insignificant role in the OCD approach.

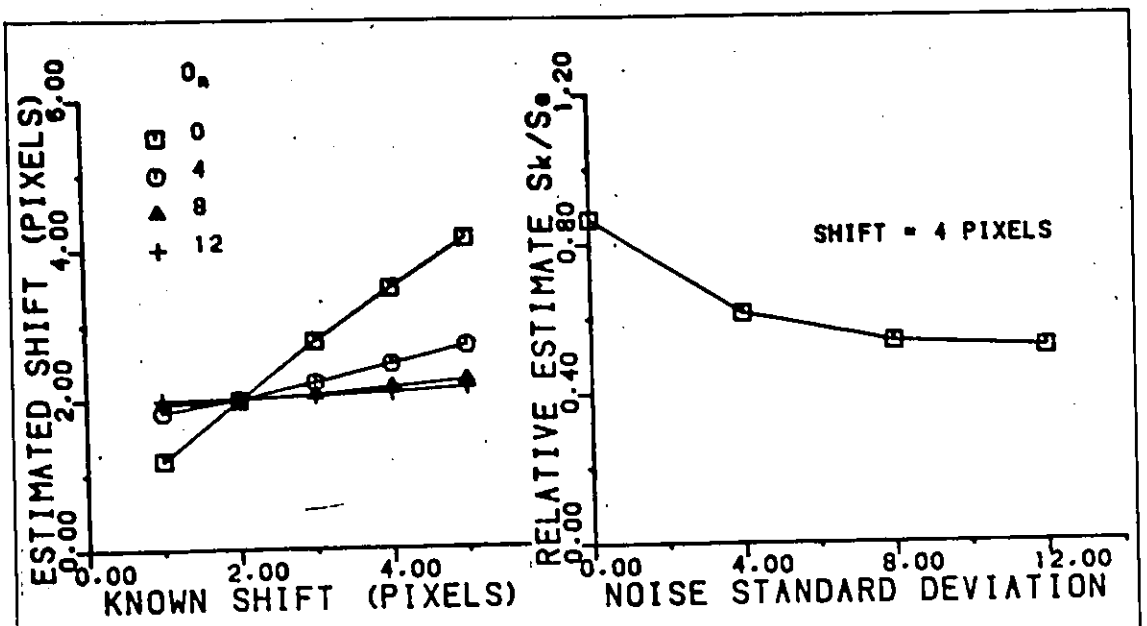


Figure 27: Effect of image noise on OCD. a) and b): The curves illustrate how image noise can greatly affect the accurary of the motion estimate.

5.4.2 Image Pre-Smoothing

By low-pass filtering the image before the use of equation 66, the performance for a large displacement can be improved and the result is shown in Figure 28 for the central block of 256x256. The smoothing process was performed by replacing eash pixel of BROWN1 by its neighbourhood average

before the calculation of the gradient and the frame difference. For example, the updated value of pixel Po using a 3x3 window can be computed as

$$P_0 = \frac{1}{8} \sum_{i=0}^{8} P_i \quad \text{for a heighbourhood} \quad \begin{matrix} P_1 & P_2 & P_3 \\ P_8 & P_0 & P_4 \\ P_7 & P_6 & P_5 \end{matrix} \tag{68}$$

Low pass filtering or smoothing reduces noise spikes, local irregularity and widens the narrow transition region of sharp edges. From the curves of Figure 28, it can be seen that the error of estimation decreases as the smoothing window size increases. The result also shows that in this case of relatively low noise, a good estimates can be obtained even with the original (1x1 smoothing) image.

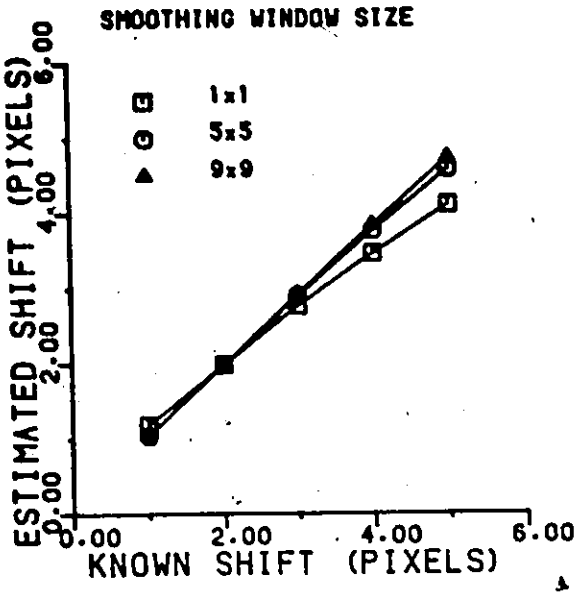


Figure 28: Effect of Smoothing Window Size on OCD. The accurary of the estimate can be improved through the use of a large smoothing window size. In this case of a low noise image, good estimate can be obtained even with the original (1x1 smoothing) image.

5.4.3 Block-size

In general, the distribution of edges and the amount of detail in an image is different from one part of the image to another. As a result, the performance of the OCD approach is influenced by the texture of the image area involved in the computation. For a small block, the local image statistics may include no edge information or may be full of details so that the estimated values of shifts differ considerably from the true values. However, for a large block the effect of the variation in local image statistics

may be reduced and a better estimate may result. In this simulation, displacement estimates are obtained with the central area of BROWN1 using a block size of 256x256, 64x64 and 8x8, and the results are shown in Figure 29.

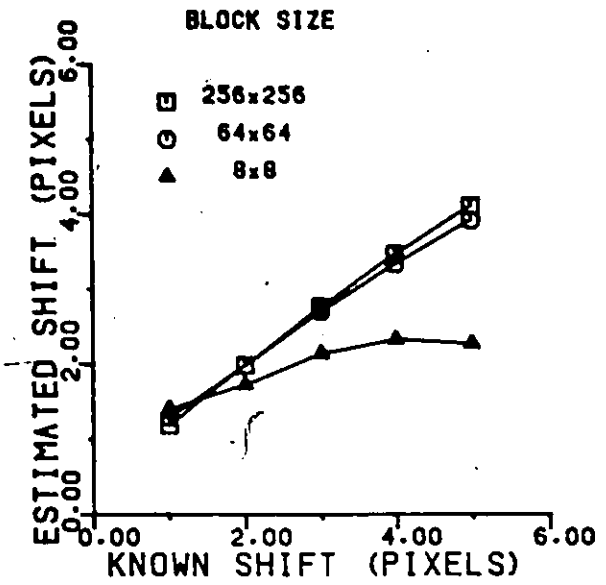


Figure 29: Effect of Block-size on OCD. The curves illustrate how the effect of the variation of local image statistics can be reduced through the use of a large block size and thus provide a better estimate.

5.5 HISTOGRAM SKEWNESS OF THE DIFFERENCE IMAGE: VESSEL OPACIFICATION

In order to study the statistics of the difference image due to vessel opacification, the various angiogram sequences described were studied and procedures similar to the previous studies were performed. The results are shown in Table

5 with measurements computed for an arbitrary central (256x256) region from the maximum opacification difference image obtainable in each sequence (Figure 30).

TABLE 5

Effect of Dye

	measure of skewness	displacement estimate
BROWN	-0.79	2.16
GREY	-2.91	7.02
WHITE	-0.46	7.29
BLACK	-2.65	6.32

The two columns listed are the histogram skewness of the difference images and the prospective estimate of motion for each of the image sequences. The result shows that GREY and BLACK have a large prospective false estimate of motion but are associated with a large negative skewness. BROWN has a smaller prospective estimate and negative skewness due to a lower contrast opacification. WHITE, however, has a large prospective displacement estimate but does not show a significant negative skewness since the central region of WHITE does not correspond to the opacification region. It must therefore be assumed that the large prospective displacement estimate is caused by the effect of image noise instead.

With regard to these observations, it is clear that there exists a limit to the amount of displacement that can reliably be detected without increasing the risk of a false detection of motion caused by vessel opacification. For example, if the amount of motion allowable in the case of BROWN is 1 pixel, it would not be possible to correctly detect patient motion without a false alarm. However, for a low contrast difference which generally exhibits opacified vessel of large diameter, the allowable motion is larger.

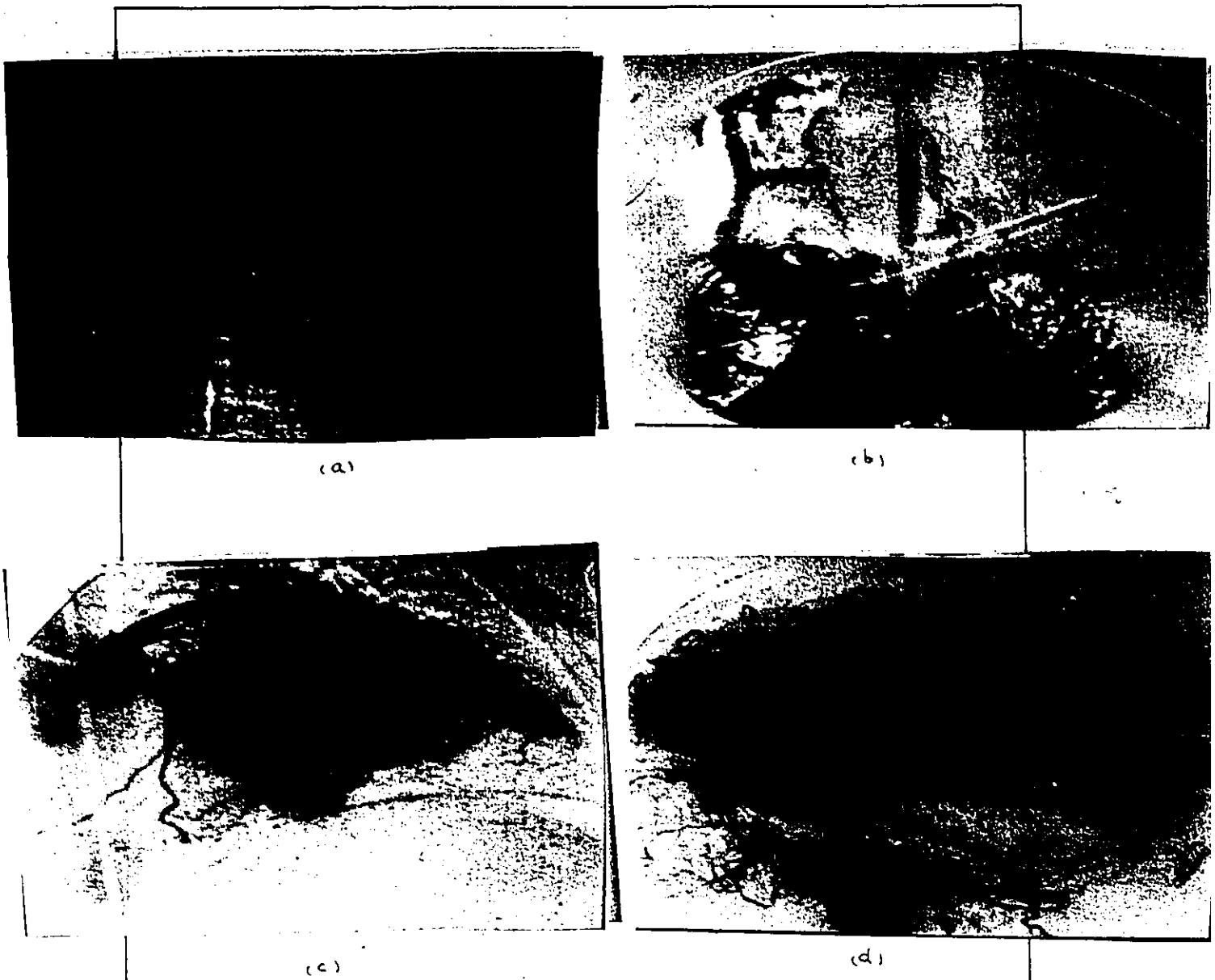


Figure 30: Intravenous angiogram performed with temporal subtraction. a) BROWN: carotid arteries in the neck of a patient, b) GREY: aortic arch examination, c) WHITE: coronary arteries supplying blood to the heart, d) BLACK: intracranial vessels within the skull.

5.6 DETECTION CRITERION

In this section, the entire BROWN sequence is analyzed in order to study the practicality of the OCD approach in a typical angiogram sequence involving motion. Image statistics were computed using the central area of 256x256 pixels and the results are shown in Table 6. The first row of the table corresponds to data collected for the difference image obtained by subtracting frame 1 from frame 2, so that the first entry in the bracket corresponds to the live image and the second entry is the mask. The motion case of frame 2 - frame 1 illustrates the unique characteristics of a large displacement with no significant negative skewness of the difference histogram. , That is, there is a large displacement estimate with no negative skewness.

TABLE 6

Case of angiogram with motion: BROWN

Frame	measure of skewness	displacement estimate
(2 - 1)	0.14	4.21
(3 - 2)	-0.20	1.51
(4 - 2)	-1.78	1.90
(5 - 2)	-1.12	2.64
(6 - 2)	-0.79	2.16
(7 - 2)	-0.73	2.49

The rest of the table corresponds to the results obtained after the reselection of a new mask (ie., frame 2), following the detection of a significant motion estimate in the preceeding subtraction. It can be seen that the displacement estimates are significantly reduced and negative skewness can be observed. In certain cases, such as the case of (5-2) and (7-2), larger displacement estimates are observed in association, in general, with a large negative skewness. This is consistent with expectation since it corresponds to changes in vessel opacification.

Chapter VI

CONCLUSION AND FUTURE RESEARCH POSSIBILITIES

This dissertation describes investigations which were made to study the problem of automated patient motion detection in DSA. A solution is proposed using a real-time remasking scheme which minimizes the effect of motion by reselecting the mask image whenever a significant motion is detected. The problem of motion estimation is analyzed from an image processing point of view and a statistical approach for real time motion detection is proposed. Instead of relying on the spatial-temporal relationship of individual pixels, the overall contrast of the difference image is explored and translational motion is adopted for an analytical study of motion. Motions caused by rotation and occulsion have not been considered. Such effects are not expected to affect the performance of the detection process since a large motion of this type will also cause an increase in the overall contrast of the difference image.

The approach is simulated on clinical angiogram images to ascertain its usefulness. The performance of the OCD approach to motion detection is studied on simulated motion generated by a translation of the image and results better than the LSE are observed. The results show that in a

noiseless situation, good results are obtained even without the use of pre-smoothing. It has been observed that large block size should be used in order to reduce the effect of the variation of local statistics. However, image noise is a limitation to the approach and image smoothing would be required in a noisy enviroment.

Of central importance to this research is the fact that it has not always been possible to clearly distinguish between changes due to the arrival of dye and those due to significant motion. To reduce the chance of a false detection of motion cause by vessel opacification, the skewness of the histogram of the difference image is studied. The results suggest that such image change is generally associated with a large skewness of the difference image histogram if the block size includes as much of the opacified area as possible. As the contrast level decreases, the correct detection of a slight motion would not be possible.

This research work contributes to a beginning of an interdisplinary work of applying image sequence analysis in medicine. Fundamental knowledge such as the principles of Digital Radiology and the analysis of image sequences formed the basis of this work which would not be complete without both. Throughout the presentation, an analytical approach was taken and a new algorithm for the detection of motion is developed. In order to improve the effectiveness of the al-

gorithm, it is expected that more a priori knowledge and modelling of the iodine opacification characteristics such as the exponential attenuation property of x-ray, the poisson property of photon detection and temporal variation of contrast flow will be required. Thus far, an analytic or numerical analysis describing the probability of detection has not been obtained. With further improvement using more a priori knowledge and with the availability of clinical data, a numerical value indicating the probabilities of detection and of false detection may be obtained.

Appendix A

ANALYSIS OF NOISE LEVEL IN DIGITAL RADIOGRAPHIC IMAGES

A derivation of the noise levels of x-ray statistical and electrical components as a functiom of signal level is given here and the following notation is used.

V_{max} = video signal voltage when the maximum useable signal current is read from the camera terget.

σ_E = the rms additive noise voltage in the video signal due to a source other than x-ray statistics (eg. video noise).

SNR = video signal to noise ratio = V_{max}/σ_E

N_o = the detected x-rays per pixel corresponding to V_{max}

α = a gain, such as video camera aperture, which maps N_o to V_{max}

N = detected x-rays in the pixel of interest,
= N exp($-\mu x$) where x is the incremental object thickness.

The video signal read from the camera is given by:

$$V = \alpha N + m \tag{A.1}$$

where m represents the video noise.

When the signal is passed through a logarithmic compression, the log signal, S, is given by:

$$S(V) = \log(V)$$
$$= S(\bar{V}) + S'(\bar{V})(V-\bar{V}) + S''(\bar{V})\frac{(V-\bar{V})^2}{2} + \ldots. \tag{A.2}$$

With the Taylor expansion, the following equations can be derived:

$$E[S(V)] = S(\bar{V}) + S''(\bar{V})\frac{\sigma_v^2}{2} + \ldots. \quad (A.3)$$

$$E^2[S(V)] = S^2(\bar{V}) + S(\bar{V})S''(\bar{V})\sigma_v^2 + \ldots. \quad (A.4)$$

$$E[S^2(V)] = S^2(\bar{V}) + [S'(\bar{V})^2 + S(\bar{V})S''(\bar{V})]\sigma_v^2 + \ldots. \quad (A.5)$$

Using the equations (A.4) and (A.5) yields

$$\begin{aligned}\sigma^2 &= E[S^2(V)]-E^2[S(V)] \\ &= [S'(\bar{V})^2] \quad + \text{higher order terms} \\ &\cong \frac{\alpha^2\bar{N}+\sigma_E^2}{\alpha^2N^2}\end{aligned} \quad (A.6)$$

since the x-ray photon counts exceeds 10^4 and N and m are independent.

It has been assumed that electronic and quantum noise are independent, and that the average value of m is zero. With suitable substitutions, σ^2 can be reexpressed as:

$$\sigma^2 = \frac{\exp(\mu x)}{N} + \frac{\exp(2\mu x)}{SNR^2} \quad (A.7)$$

The first term is the noise contributed by x-ray statistics and the second term is the contribution due to video noise.

Appendix B

DERIVATION OF THE LEAST SQUARE ESTIMATE

With a set of points $\{X(i),Y(i),T(i)\}$, the problem of fitting a plane to the data can be solved through the use of linear regression. If T(i) of the data is estimated by

$$T_e(i) = -\{aX(i)+bY(i)\}$$

where a and b are constants, the error e(i) in estimating T(i) for a given i is

$$T(i)-T_e(i) = e(i)$$

To provide the best fit, the constants a and b are determined so that the estimated errors are minimized according to the criterion:

$$\sum_{i=1}^{n} e^2(i) \quad \text{is minimized}$$

In other words, the constants a and b are chosen such that

$$\sum_{i=1}^{n} \{T(i)+aX(i)+bY(i)\}^2 = \min$$

Since a necessary condition for a relative minimum is the vanishing of the partial derivatives with respect to a and b,

$$2 \sum \{T(i)+aX(i)+bY(i)\} \{X(i)\} = 0$$

$$2 \sum \{T(i)+aX(i)+bY(i)\} \{Y(i)\} = 0$$

Solving the two equations simultanously yields

$$a = \frac{(\Sigma TX)(\Sigma Y^2) - (\Sigma XY)(\Sigma TY)}{(\Sigma XY)^2 - (\Sigma Y^2)(\Sigma X^2)}$$

and

$$b = \frac{(\Sigma TY)(\Sigma X^2) - (\Sigma XY)(\Sigma TX)}{(\Sigma XY)^2 - (\Sigma Y^2)(\Sigma X^2)}$$

where the subscript have been omitted for clearity. Therefore, with a given set of points (X,Y,T), the plane of best fit can be defined by obtaining the constants a and b using the above solution.

REFERENCES

1. P.E. Anuta "Spatial Registration of Multispectral and Multitemporal Digital Imagery using FFT Techniques," IEEE Trans. on Geoscience Electronics Vol. GE-8, No.4, pp.353-68 Oct.1970.

2. A.L. Baert, G. Wilms (et al.) "Intravenous Digital Subtraction Angiography," Europ. J. Radio. 1:97-103, 1981.

3. D.I. Baniel, H.F. Silverman "A Class of Algorithm for Digital Image Registration," IEEE Trans. on Computer Vol.C-21, pp.179-86, Feb.1972.

4. S.T. Barnard W.B. Thompson "Disparity Analysis of Images," IEEE Trans. on Pattern Analysis and Machine Intelligence Vol.PAMI-2, No.4, pp.333-40 ,July 1980.

5. W.R. Brody "Hybrid Subtraction for improved Arteriography," Radiology Vol.141:828-31, Dec. 1981.

6. W.R. Brody "Digital Subtraction Angiography," IEEE Trans. on Nuclear Science Vol.NS-29, No.3, pp.1176-80, June 1982.

7. H. Bruggemann "Temporal filtering using Pixel Incrementing," SMPTE Journal pp.686-94, August 1981.

8. T.F. Budinger "Physical and Medical attributes of six contemporary non-invasive imaging techniques," SPIE Vol.314, pp.9-17, Sept 1981.

9. C. Cafforio, F. Rocca "Methods for measuring Small Displacements of TV images," IEEE Trans. on Information Theory Vol.IT-22, No.5, pp.573-79, Sept. 1976.

10. C.K. Chow, S.K. Hital (et al.) "X-ray Image Subtraction by Digital Means," IBM J. Res. Develop. Vol.17, pp.206-18, 1973.

11. L.S. Davis "A survey of Edge Detection Technique," Computer Graphics and Image Processing Vol.4, pp.248-70, 1975.

12. D.P. DeBrust "Digital X-ray Imaging: The Basics," Radiologic Technology Vol.54, No.4, pp.280-86

13. "Digital Radiography: An assessment of its impact on radiological practice," General Electric Company, 1980.

14. "Digital Vascular Imaging," Philips Company

15. P.H.C. Engels, J.W. Ludwig (et al.) "Left Ventricular Evaluation by Digital Video Subtraction Angiocardiography" Radiology 144:471-74 August 1982

16. C.L. Fennema, W.B. Thompson "Velocity Determination in Scene Containing Several Moving Objects," Computer Graphics and Image Processing Vol.9, pp.301-15, 1979.

17. R.C. Gonzalez, P. Wintz Digital Image Processing, Addison-Wesley Publication Co., Massachusetts, 1977.

18. R.G. Gould "Time Subtraction digital fluoroscopy with a dual memory system," IEEE Trans. on Nuclear Science Vol.NS-29, No.3, pp.1186-90, June 1982.

19. V. Graefe "On the representation of moving objects in real-time computer vision systems," SPIE Vol.432, pp.129-35

20. U.L. Haass "A visual surveillance system for tracking of moving objects in industrial workroom environment," Proc. 6th International Conference on Pattern Recognition pp.757-59, 1982.

21. U.L. Haass, H.B. Kuntse (et al.) "A surveillance system for obstacle recognition and collision avoidance control in robot environment,"

22. D.P. Harrington, L.M. boxt (et al.) "Digital Subtraction Angiography: Overview of Technical Principles," American Journal of Radiology Vol.139:781-86, Oct.1982.

23. G. Hass, V. Graefe "Locating fast moving objects in TV. images in the presence of motion blur," SPIE: Application in Digital Image Processing pp.440-46, 1983.

24. S.M. Haynes, R. Jain "Time Varying edge detection," Proc. 6th International Conference on Pattern Recognition pp.754-56, 1982.

25. S.M. Haynes, R. Jain "Detection of Moving Edges," Computer Vision, Grephics and Image Processing Vol.21, pp.345-67, 1983.

26. R.V. Hogg, E.A. Tanis Probability and Statistical Inference Macmillan Pub. Co., New York, 1977.

27. K.H. Hohne (ed.) Digital Image Processing in Medicine SPringer-Verlag, Berlin Heidelberg, 1981.

28. B.K.P. Horn, B.G. Schunch "Determining optical flow," Artifical Intelligence Vol.17, pp.185-203, 1981.

29. T.S. Huang (ed.) Image Sequence Analysis Springer-Verlag, New York, 1981.

30. T.S. Huang (ed.) Image Sequence Processing and Dynamic Scene Analysis Springer-Verlag, Berlin Heidelberg, 1983. B.R. Hunta, T.M. Cannon "Non-Stationary Assumption for Gaussian Models of Images," IEEE Trans. on System, MAn and Cybernetics pp.876-82, Dec. 1976.

31. R. Jain, S.N. Jayaramamurthy "Time-Varing Image Analysis" pp.525-29, 1982.

32. R. Jain, H.H. Nagel "On the analysis of Accumulative Difference Pictures from image sequences of Real World Scene," IEEE Trans. on Pattern Analysis and Machine Intelligence Vol.PAMI-1, No.2, pp.206-14, April 1978.

33. J.K. Kearney, W.B. Thompson (et al.) "Gradient Based Estimation of disparity ," Proc. of the Pattern Recognition and Image Processing pp.246-51, June 1982.

34. G.S. Keyes, N.J. Pelc (et al.) "Digital Flurography: A Technology Update," General Electric Company 1981.

35. R.A. Kruger "A Method for Time domain filtering using computerized fluoroscopy" The 67-th Scientific Assembly and Annual Meeting

36. R.A. Kruger, F.J. Miller (et al.) "Digital Subtraction Angiography Using a Temporal Bandpass Filter: Associated Patient Motion Properties," Radiology Vol.145:315-20, Nov.1982.

37. R.A. Kruger, C.A. Mistretta "Physical and Technical considerations of Computerized Fluoroscopy Difference Imaging," IEEE Trans. on Nuclear Science Vol.NS-28, No. 1, pp.205-12, Feb.1981.

38. R.A. Kruger, Pingyu Liu (et al.) "Time domain filtering using computerized fluoroscopy: Intravenous Angiography application," SPIE Vol.314:319-26, Sept. 1981.

39. R.A. Kruger, Pingyu Liu "Digital Angiography Using a Matched FIlter," IEEE Trans. on Medical Imaging Vol.MI-1, No.1, pp.16-21, July 1982.

40. R.A. Kruger "Vascular Tracing with Computerized Fluoroscopy" Patent Disclosure April 15, 1981.

41. J.O. Limb, J.A. Murphy "Estimating the velocity of moving images in TV. signals," Computer Graphics and Image Processing Vol.4, pp.311-27, 1975.

42. J.O. Limb, J.A. Murphy "Measuring te speed of moving objects from TV. signals," IEEE Trans. on COmmunication pp.474-78, April 1975.

43. L. Maisel Probability, Statistics and Random processes Simon and Schuster Technical Outline, New York, 1971.

44. W.N. Martin, J.K. Aggarwal "Dynamic Scene Analysis," Computer Graphics and Image Processing Vol.7, pp.356-74, 1978.

45. C.A. Mistretta "Current Pratice and future directions in digital subtraction angiography," SPIE Vol.314 pp.18-23, Sept. 1981.

46. C.A. Mistretta, A.B. Crummy (et al.) "Digital Angiography: A perspective," Radiology 139:273-76, May 1981.

47. C.A. Mistretta, R.A. Kruger (et al.) "Digital Vascular Imaging,"

48. H.H. Nagel "Analysis Techniques for Image Sequences," Proc. 4th International Joint Conference on Pattern Recognition pp.186-211, 1978.

49. H.H. Nagel "Constraints for the estimation of displacement vector fields from Image Sequences," Proc. 8th International Joint Conference on Artificial Intelligence pp.945-51, 1983.

50. A.N. Netravali, J.O. Limb "Picture Coding: A review," Proceedings of IEEE Vol.68, No.3, pp.366-407, March 1980.

51. A.N. Netravali, J.O. Robbins "Motion Compensated TV. coding" Bell System Technical Journal Vol.58, No.3, March 1979.

52. H. Oung, A.M. Smith "Real time motion detection in digital subtractive angiography," International Symposium on Medical Images and ICON July 1984.

53. D.P. Panda, T. Dubitski "Statistical Analysis of some edge operators," Computer Graphics and Image Processing Vol.11, pp.313-48, 1979.

54. L.J. Pinson "High resolution images registration by thresholded difference," SPIE: Application of Digital Image Processing Vol.359, pp.178-85, 1982.

55. A. Popoulis Probability, Random Variables and Stochastic Processes McGraw-Hill, 1965.

56. H.J. Potel, D.E. Gustafson "Motion Correction for Digital Subtraction Angiography," Proc. 5th IEEE EMBS pp.166-69, Sept. 1983.

57. W.K. Pratt Digital Image Processing Wiley, New York, 1978.

58. S.J. Riederer, F.A. DiBianca (et al.) "Performance characteristics of a digital fluorographic system," SPIE: Application of Optical Instrumentation in Medicine IX Vol.273, pp.88-95, 1981

59. S.J. Riederer, A.L. Hall (et al.) "The technical characteristics of matched filtering in digital subtraction angiography," Medical Physics Vol.10, No.2, pp.209-17, Mar./April 1983.

60. S.J. Riederer, R.A. Kruger "Intravenous Digital Subtraction: A summary of recent developments," Radiology Vol.147:633-38, June 1983.

61. A. Rosenfeld, A.C. Kak Digital Picture Processing Academic Press, New York, 1982.

62. C.G. Shaw, D.L. Ergun (et al.) "Intravenous Angiography using computerized fluoroscopy," IEEE Trans. on Nuclear Science Vol.NS-27, No.3, pp.1042-46, June 1980.

63. A.M. Smith, M. Goldberg (et al.) "Digital Fluoroscopic Processor," Project Report, U. of Ottawa, 1982.

64. M. Svedlow, C.D. McGillem (et al.) "Image Registration: A similarity measure and preprocessing method comparisons," IEEE Trans. on Aerospace and Electronic Systems Vol.AES-14, No.1, pp.141-49, Jan.1978.

65. W.B. Thompson "Combining motion and contrast for segmentation," IEEE Trans. on Pattern Analysis and MAchine Intelligence pp.543-49, Nov. 1980.

66. W.B. Thompson, S.T. Barnard "Low level estimation and interpretation of Visual Motion," Computer pp.20-28, Aug. 1981.

67. M.S. Van Lysel, W.C. Zarnstorff (et al.) "Real time digital video recording system," SPIE Vol.314, pp.389-95, Sept. 1981.

68. J.G. Webster (ed.) Medical Instrumentation: Application and design Houghton Mifflin Company, Boston, 1978.