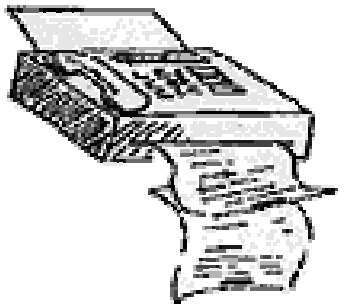
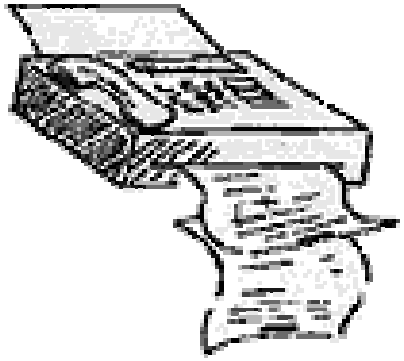
FACSIMILE: CODING AND TRANSMISSION OF BILEVEL IMAGES



Fernando Pereira

Instituto Superior Técnico



Facsimile: Objective

Efficient representation of bilevel images for transmission using telephone and data networks.

History of Facsimile (1)

- **1843 – First facsimile patent (England, nº 9745) registered by Mr. Alexander Bain –** ***telephone has not been invented until 1876*** **!**
- **1843 - ? - Main problemas to solve at that time were power sources, scanning, synchronization, transmission channel (telegraph line).**
- **1865 – First commercial between Lion and Paris.**
- **1876 – Telephone emerges ...**
- **1911 – First modulator for facsimile transmission over the telephone line.**
- **1900 ... – Along XX century many technological advances have been made related to the various parts of a facsimile system.**

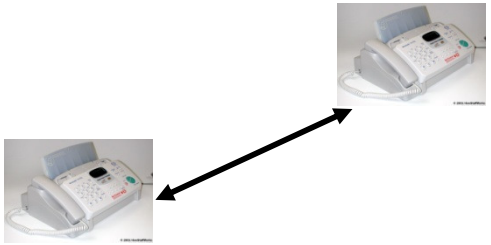
History of Facsimile (2)

- **1969 – First digital fax appears ...**
- **1974 and 1976 – Standards for analogue fax - *groups 1 and 2* - appear.**
- **1980 – Group 3 digital fax standard appears allowing the quick spreading of this type of terminals.**
- **1984 – Group 4 digital fax standards appears targetting transmission over data networks.**
- **1991 – Further improvements on group 3 facsimile; group 3 faxs have 99.7 % of the market with more than 20 million terminals.**
- **199x – Internet takes the fax market share ...**

Standard Facsimile Equipment (Recommendation ITU-T T.0)

- ✶ Faxs using telephone network transmission:
  - Analogue
    - **GROUP 1** – Uses double band amplitude modulation without any (analogue) compression of the transmission bandwidth; the transmission of an A4 page takes about 6 minutes for a resolution of 3.85 linhas/mm (recommendation T.2)
  - Analogue
    - **GROUP 2** – Uses bandwidth compression techniques (vestigial side band) to obtain a transmission time of about 3 minutes for an A4 page with a resolution of 3.85 linhas/mm; any processing for redundancy reduction is excluded (recommendation T.3)
  - Digital
    - **GROUP 3** – Uses redundancy reduction digital processing techniques before modulation; the transmission of an A4 page takes about 1 minute for a resolution of 3.85 linhas/mm (recommendation T.4)
- ✶ Faxs using data network transmission:
  - Digital
    - **GROUP 4** – Uses redundancy reduction digital processing techniques and operates over public data networks, which provide a virtually error free transmission (recommendations T.5 and T.6)

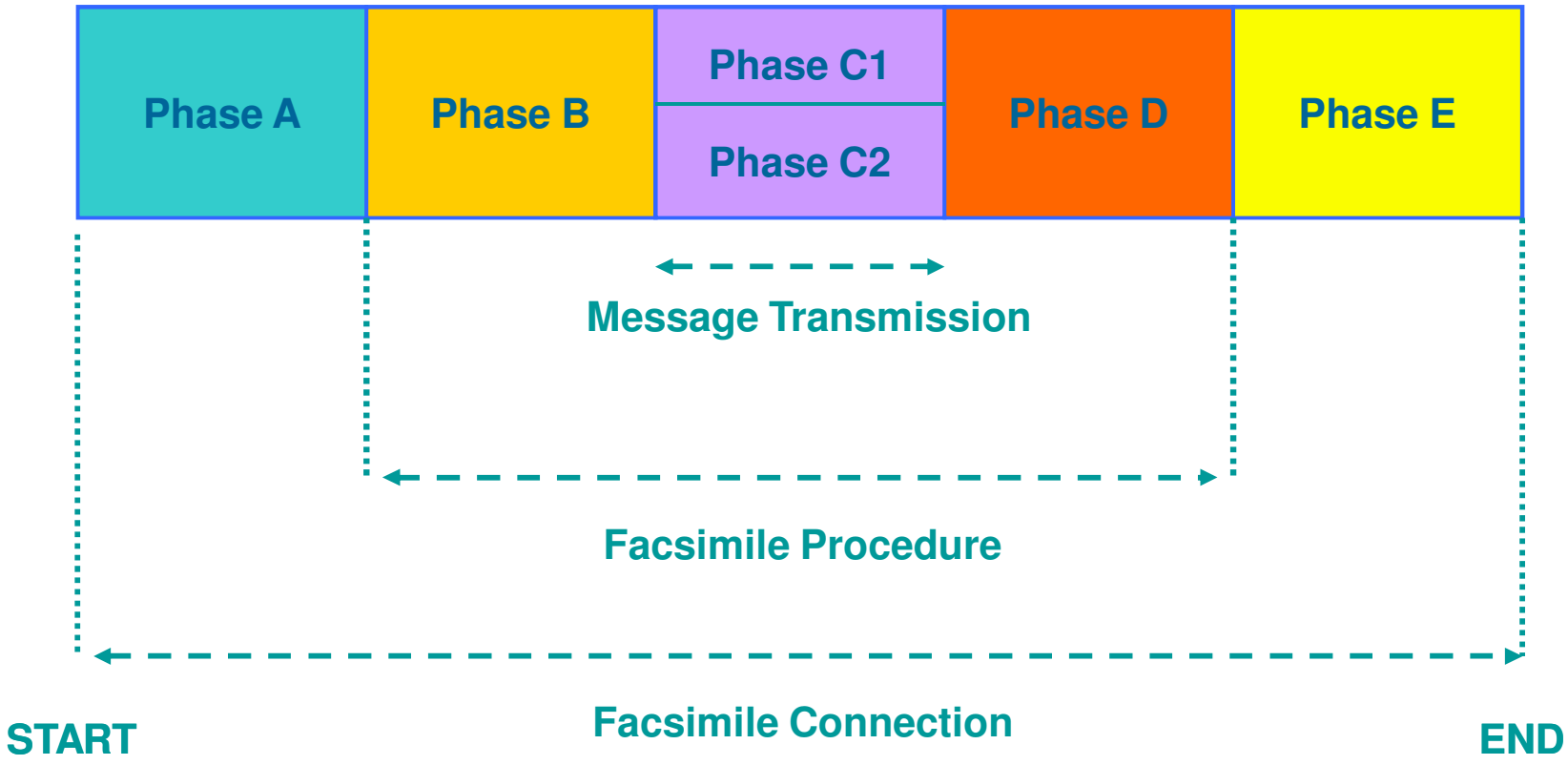
Communication Protocol



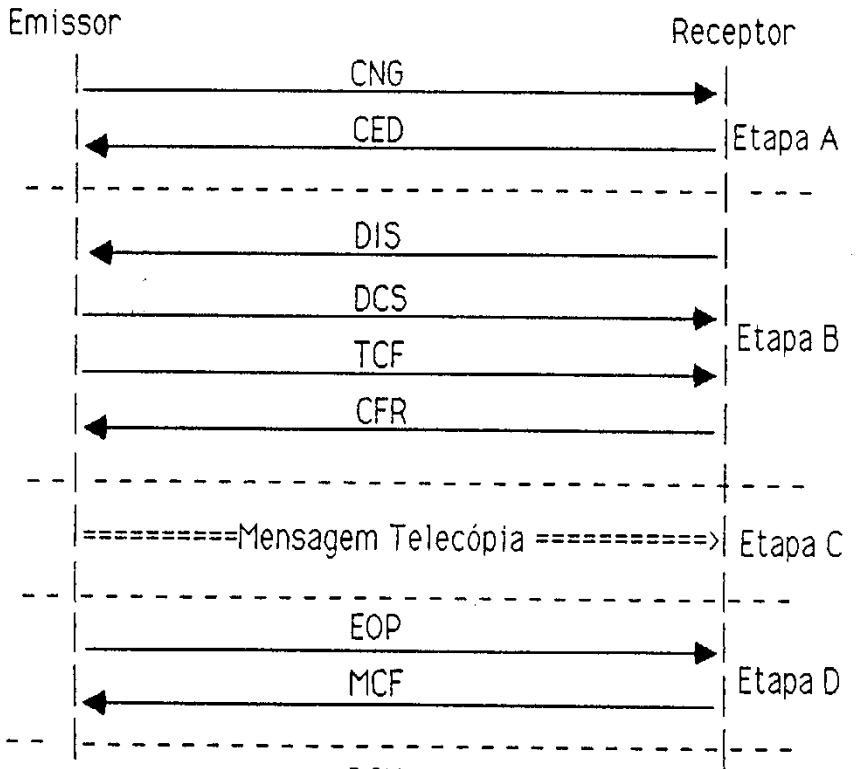
Recommendation T.30 specifies the protocol for the transmission of facsimile documents over the telephone network.

- **Phase A – Call Setup:** the fax connection is established using a specified protocol based on sinusoidal tones.
- **Phase B – Pre-Message Procedure:** the 2 faxs exchange their capabilities to agree on operational conditions; the calling fax is always the one leading.
- **Phase C – Message Transmission**: the image information is sent using the operational parameters previously agreed.
- **Phase D – Post-Message Procedure:** the 'good' reception is confirmed; more pages may be sent or the connection is finished.
- **Phase E – Call Release:** Both fax machines disconnect from the telephone line.

Phases of a Facsimile Call



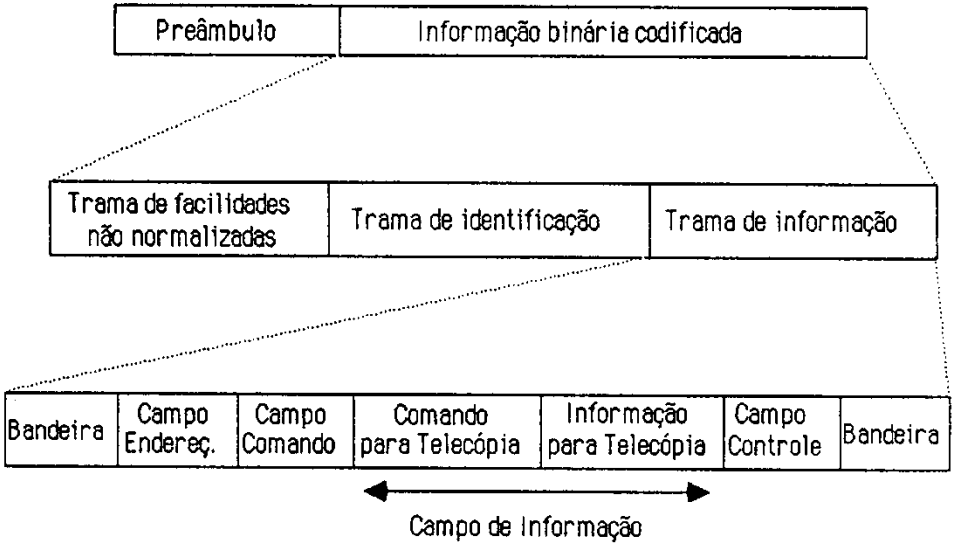
Group 3 Protocol



- **CNG – Calling signal - every 3.5 s a 1100 Hz sinosoid 0.5 s long is sent.**
- **CED - Answering signal - 2100 Hz sinosoid during 2.6 to 4 s.**
- **DIS - Digital Identification Signal – caracterizes the receiving terminal in terms of standard features.**
- **DCS - Digital Command Signal – determines the connection characteristics based on the sending and receiving terminals features.**
- **TCF - Training Check – training sequence is sent to analyise the line and determine the transmission rate to use without too many errors; consists in a sequence of 0s during 1.5 s.**
- **CFR - Confirmation to Receive – confirms the preliminary procedures and determines the starting of the message transmission phase**
- **EOP - End-of-Procedure – indicates the end of the transmission of one image; if there is no need to send more images, the connection will be disconnected (after confirmation).**
- **MCF - Message Confirmation – confirms the reception of one image and the availability to receive more.**
- **DCN - Disconnect – disconnecting ...**

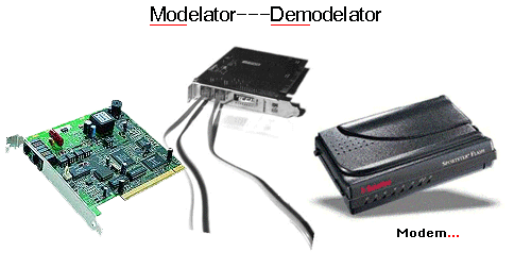
Group 3 Protocol

For all phases of the communication protocol, with the exception of the message transmission and call setup, HDLC (*High-Level Data Link Control*) frames are used.



- **Basic rules of this protocol are:**
  - **Optional frames must always be acompanied by a mandatory frame transmitted as last.**
  - **When receiving optional frames that it is not able to recognize, a terminal must discard them using only the mandatory frames received.**
  - **HDLC frames always use bit stuffing with the exception of the delimitation flags.**

Group 3 Modems



- A fax modem has the task to take digital picture information and transform (modulate) it into a convenient format to be given to the transmission channel, notably in terms of bandwidth, frequency range, etc.
- The mandatory modems for group 3 are the V.27 ter modem for the transmission of the picture information at 4.8 or 2.4 kbit/s and the V.21 modem for the initial signaling at 300 bit/s.
- Group 3 faxs automatically test the line conditions using a training sequence.
- The transmission bitrate for the picture information is the highest bitrate that can be used by both fax in presence, guaranteeing minimum transmission conditions.

Group 3 Modem Characteristics

Bitrate (bit/s)	Baud rate (baud)	Bit/symbol	Modem type	Carrier frequency	Bandwidth (Hz)
14400	2400	6	V.17	1800	550-3050
12000	2400	5	V.17	1800	550-3050
9600	2400	4	V.29	1700	450-2950
7200	2400	3	V.29	1700	450-2950
4800	1600	3	V.27ter	1800	950-2650
2400	1200	2	V.27ter	1800	1150-2450

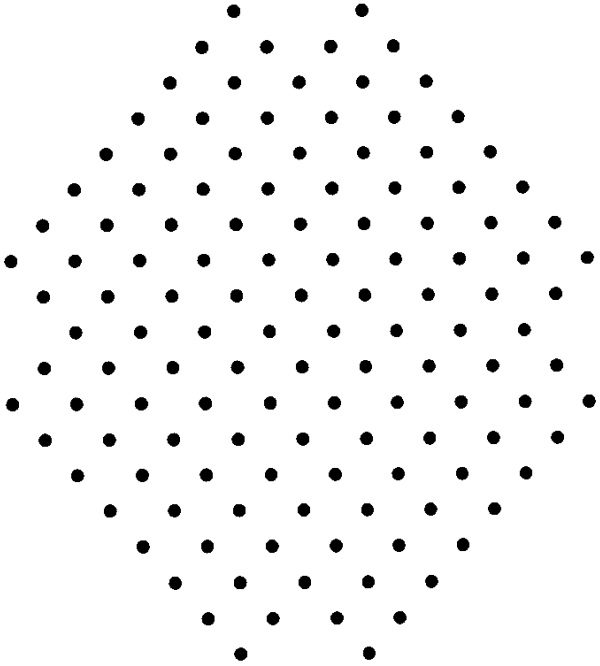
Bitrate

Bandwidth

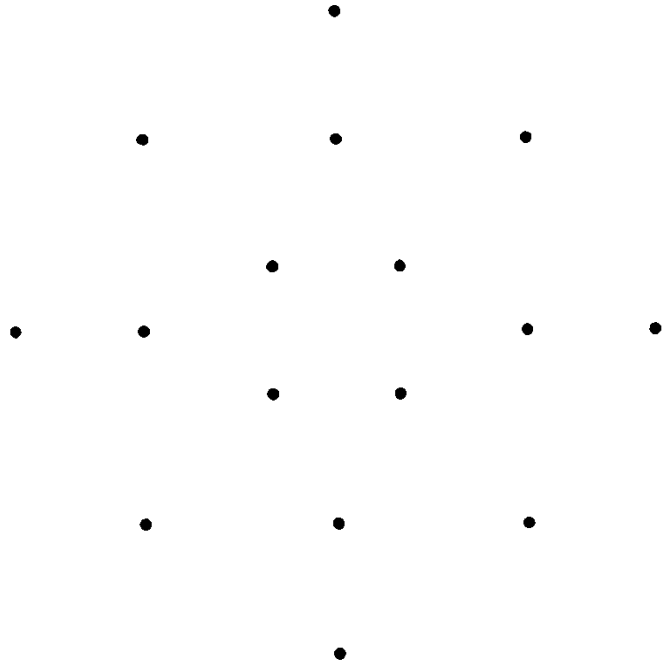
Modulation

Corresponds to the telephone channel

Modem Constelations



V.17



V.29

Group 4 Facsimile

Group 4 facsimiles operate over data networks, virtually error free, since error control protocols are present to 'clean' the connection from errors.

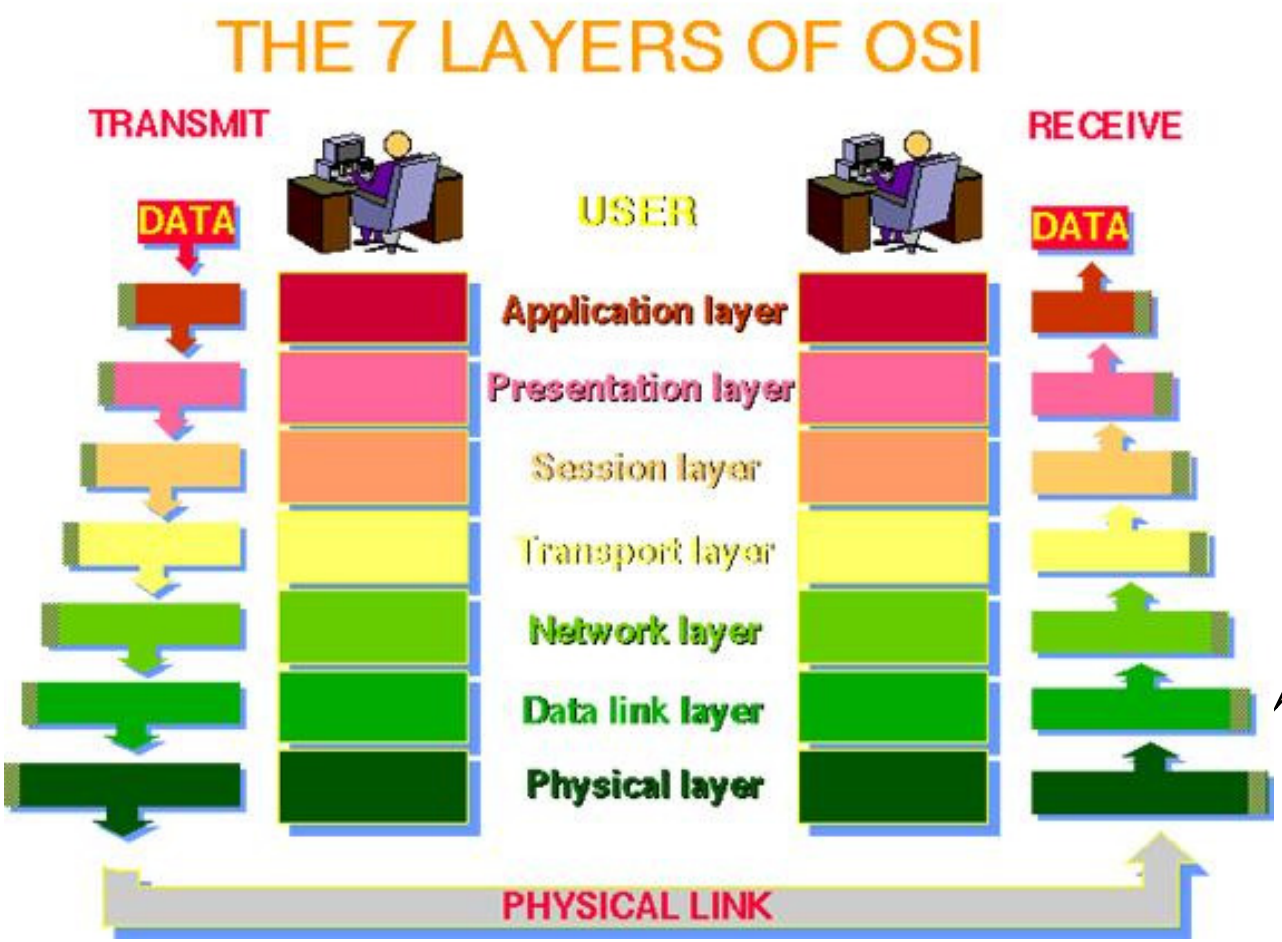
Group 4 facsimiles work as I/O terminals in remote terminals/computers.

Example group 4 facsmile applications:

- **Email – the data network is used to exchange 'mail'.**
- **Storage and retrieval – documents may be stored in a computer and accessed from a remote fax.**
- **Text and image integration – the fax terminal may digitize images that the computer processes and integrates, and later the same fax transmits.**
- **Character recognition – digitized documents may be stored after character recognition with specific purposes.**

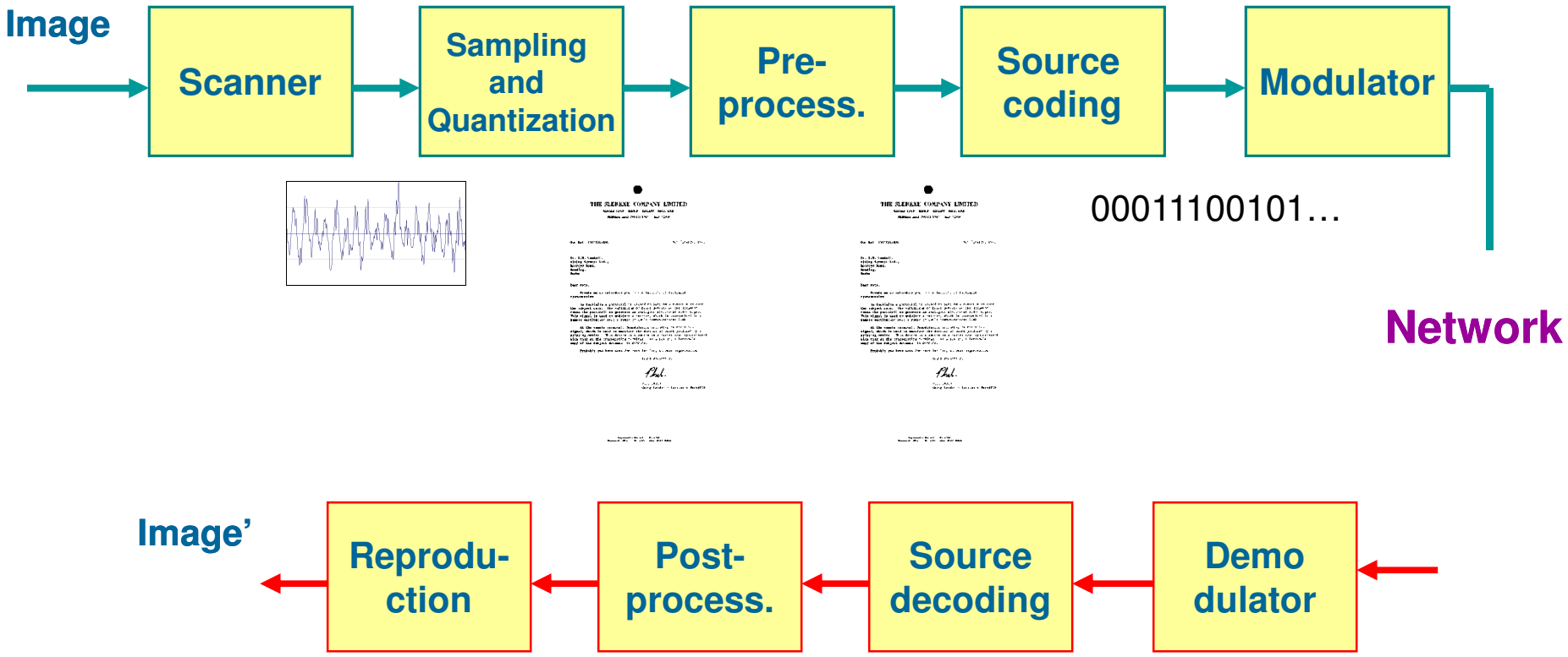
Group 4 terminals communication is assured through the OSI Model which guarantees the connection of any 2 terminals using a data network.

Group 4 Facsimiles and the OSI Model



Error detection and correction capabilities

Digital Facsimile Architecture

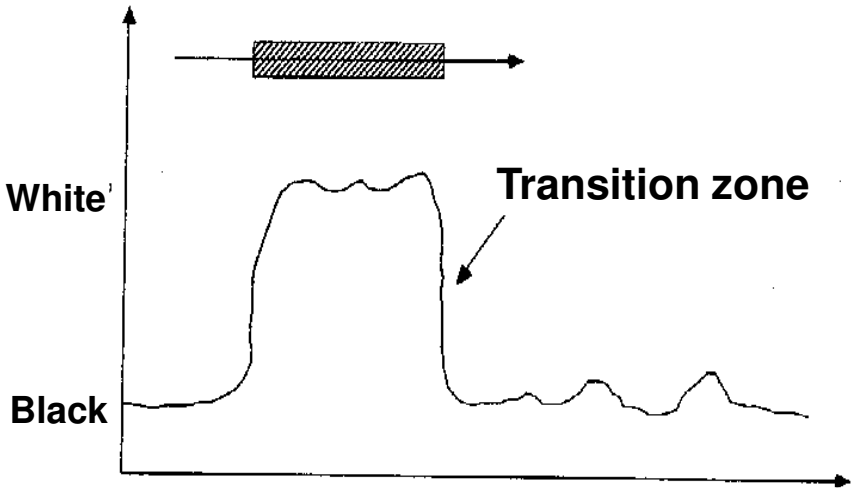


Digitization of the Image Signal

Sampling and quantization allows to obtain a digital signal from the analog output of the scanner; these processes preceed the source coding phase.

Quantization methods may be evaluated in terms of:

- Subjective quality of the associated bilevel image
- Compression factor obtained after coding
- Complexity of the quantization algorithm
- Robustness of the quantization algorithm against difficulties such as low constrast, 'recycled paper', luminance variations

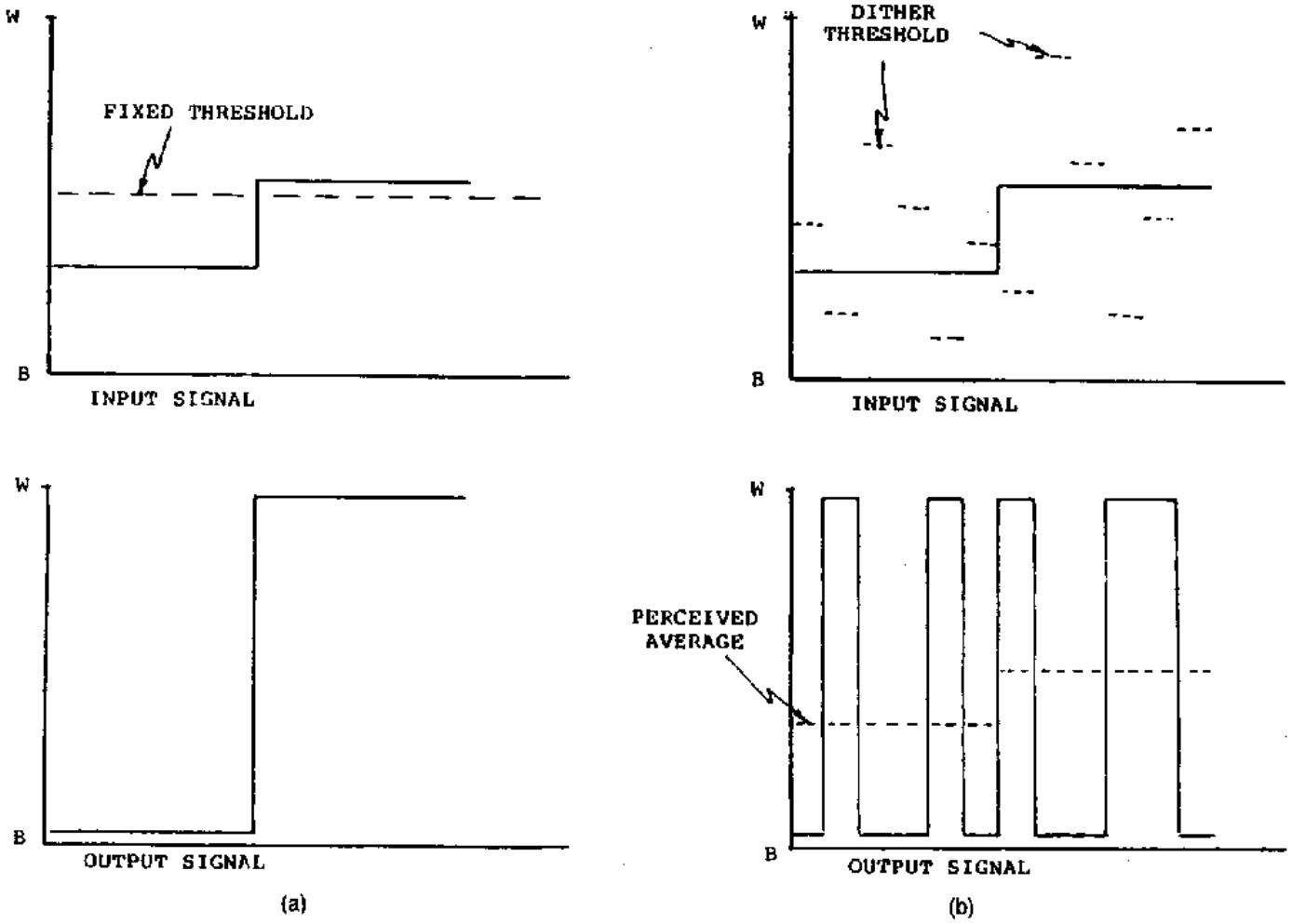


Basic Quantization Techniques

- **Fixed threshold quantization**
  - The fixed threshold depends on the gray level histogram for the signal to be quantized, which is typically the midpoint between the black and white peaks.
  - The threshold may be valid for the whole image (rigid) or just part of it (dynamic).
  - This is an acceptable quantization method for highly contrasted images but it may cause distortions for less constrasted images or when there are variations in terms of illumination or paper reflectance.

- **Variable threshold quantization (dithering)**
  - This process substantially improves the subjective quality of gray level images by allowing the threshold to uniformly vary in the full gray level range.
  - With this process, the average (black and white) luminance value in a gray zone is close to the real (gray) luminance value.

Basic Quantization Techniques: Examples



Pre-Processing for Noise Reduction (1)

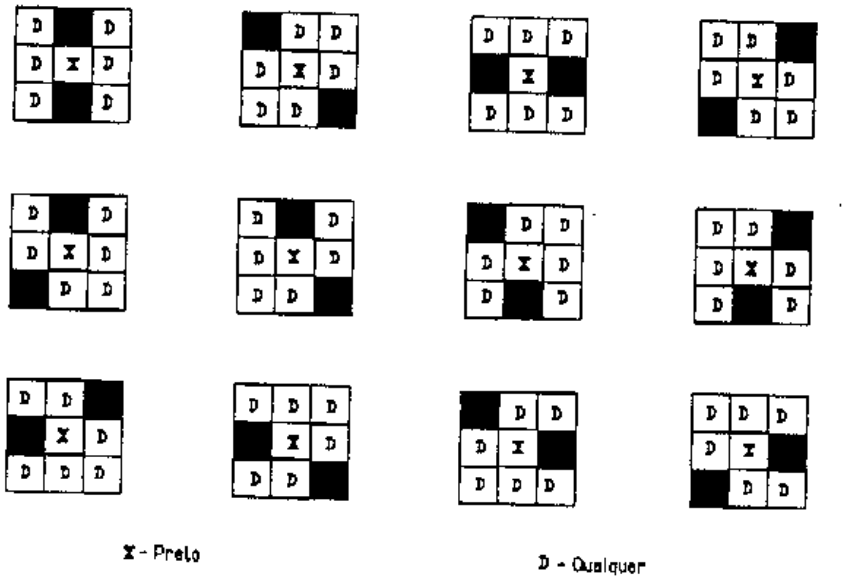
The transmission of images with 'bad quality', e.g. black dots, leads to the reduction of the compression factors and the corresponding increase of the transmission time since the spatial redundancy in the image is decreased.

Noise reduction pre-processing may 'improve' the image, making it 'cleaner', subjectively more pleasant, and allowing to reach higher compression factors.

Pre-processing may be applied to the multilevel signal at the scanner output or to bilevel signal after quantization. While the bilevel pre-processing is typically simpler, it does not allow to eliminate certain types of distortion since part of the information has already been lost in the quantization process.

Pre-Processing for Noise Reduction (2)

- **Majority processing** – The resulting value for the pixel in question is determined by the majority value for the pixels in its neighborhood.
- **Selective majority processing** – The resulting value for the pixel in question is determined by the majority value for the pixels in its neighborhood unless specific pixel configurations are present, e.g. to avoid eliminating thin lines.



Digital Image Coding

- **LOSSLESS (exact) CODING – The image is coded preserving all the information present in the digital image; this means the original and decoded images are mathematically the same.**

- **LOSSY CODING – The image is coded without preserving all the information present in the digital image; this means the original and decoder images are mathematically different although they may still be subjectively the same (transparent coding).**

Lossless coding may use pre-processing technique provided that they are reversible or applied before the signal which is taken as the original to code.

Digital Coding of Bilevel Images

GROUP 3 FAX

- **MODIFIED HUFFMAN METHOD (MHM) – Unidimensional coding method based on the coding of the lenght of alternate black and white pixel runs using Huffman coding.**

GROUP 4 FAX (also Group 3 options)

- **MODIFIED READ METHOD (MRM) – Bidimensional coding method based on the coding of the variations of the positions of tone transition pixels (black-white or white-black) in relation to the previous line; unidimensional coding may be used every *k* lines.**
- **MODIFIED-MODIFIED READ METHOD (MMRM) – Similar to MRM but without periodic unidimensional coding.**

What is a Bilevel Image ?

THE SLEREXE COMPANY LIMITED

SAPORS LANE - BOOLE - DORSET - BH 25 8 ER

TELEPHONE BOOLE (945 13) 51617 - TELEX 123456

Our Ref. 350/PJC/EAC 18th January, 1972.

Dr. P.N. Cundall,
Mining Surveys Ltd.,
Holroyd Road,
Reading,
Berks.

Dear Pete,

Permit me to introduce you to the facility of facsimile transmission.

In facsimile a photocell is caused to perform a raster scan over the subject copy. The variations of print density on the document cause the photocell to generate an analogous electrical video signal. This signal is used to modulate a carrier, which is transmitted to a remote destination over a radio or cable communications link.

At the remote terminal, demodulation reconstructs the video signal, which is used to modulate the density of print produced by a printing device. This device is scanning in a raster scan synchronised with that at the transmitting terminal. As a result, a facsimile copy of the subject document is produced.

Probably you have uses for this facility in your organisation.

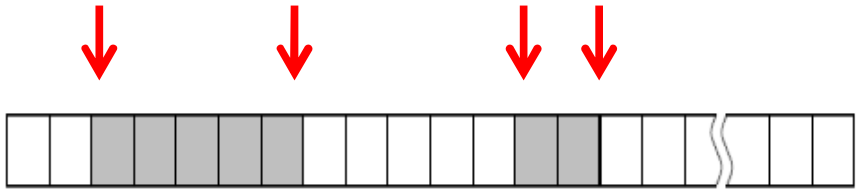
Yours sincerely,

Phil.

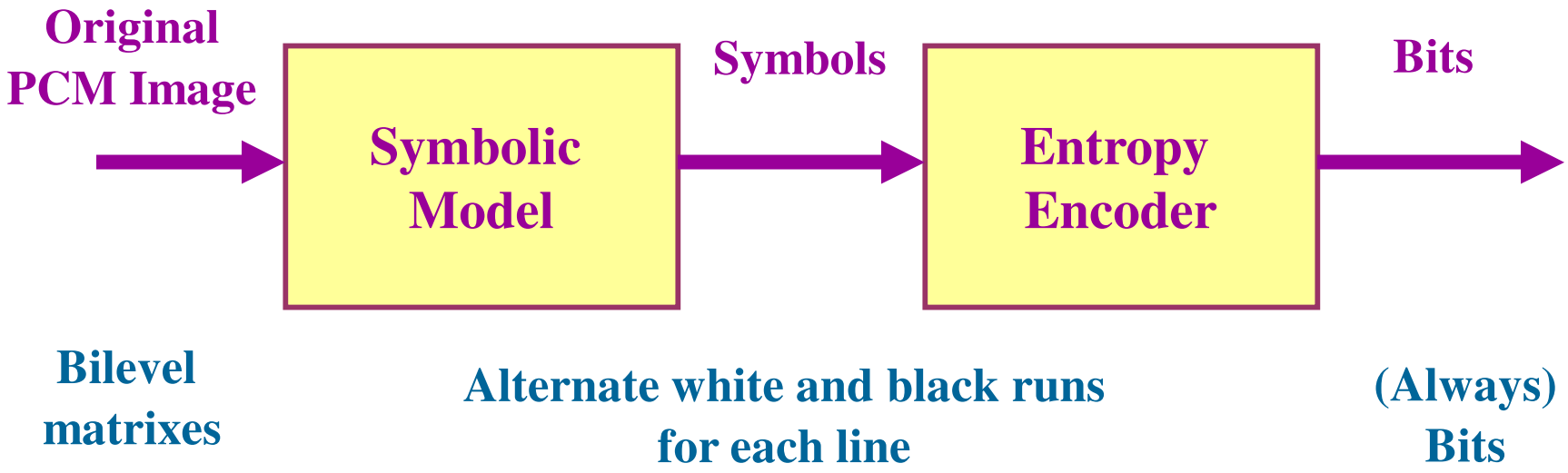
P.J. CROSS
Group Leader - Facsimile Research

Registered in England: No. 2038
Registered Office: 60 Vicars Lane, Ilford, Essex.

A bilevel image is basically a set of white-black and black-white transitions/frontiers.

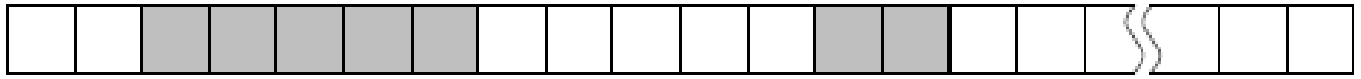


MHM Facsimile Coding



MHM Model: A facsimile image is represented as a sequence of independent lines with each line represented as an alternate sequence of white and black runs; to keep synchronism, the first run in a line is always white.

Modified Huffman Method (MHM): The Symbols



Run lengths	2B	5P	5B	2P	585B=599
MHM Code	0111	0011	1100	11	(576) 01101000 (9) 10100
Number of bits	4	4	4	2	8 + 5 = 27 bits

Compression factor: 599/27 = 22.19

- **MHM coding is based on the (indirect) representation of the *Black-White* and *White-Black* frontiers along a fax line.**
- **Each line is represented as an alternate sequence of white and black runs.**
- **For tone synchronism, first run is always white; an EOL codeword (End-Of-Line) signals the end of a line.**

Information Theory: Source Entropy

Information Theory states that there is a lower limit for the average number of bits per symbol when coding *m* symbols from a source of information, each one with probability p_i. This limit is given by the source entropy:

$$H = \Sigma\, p_i \log_2 (1/p_i) \quad \text{bit/symbol}$$

- **The source entropy:**
  - **Measures the average amount of information carried by each symbol output by the source**
  - **Is a convex function of the probabilities p_i**
  - **Takes its maximum value when all symbols are the same probability (all p_i are the same)**
  - **Takes a maximum value of $\log_2$ m bit/symbol**

Information Theory does not indicate how to obtain a code with this coding efficiency but there are methods which allow to obtain codes with an efficiency as close as desired to the entropy efficiency.

Entropy Coding

Entropy coding allows encoding into bits the symbols issued by a source taking into account their statistical distribution.

Entropy coding:

(+) Increases the final compression efficiency

(+) Does not degrade the coded signal, this means it is lossless

(-) Produces a highly time varying bitstream

(-) Increases the sensibility to transmission errors

(*) Provides compression in statistical terms, not necessarily symbol by symbol

Variable Lenght Coding (VLC)

In VLC, a codeword is attributed to each symbol which may have a different lenght. Compression is obtained by using shorter codewords for the most frequent symbols and vice-versa.

- **Codes may have the following characteristics:**
  - **Uniquely decodable – There must exist a single way to decode any sequence of VLC codes.**
  - **Instantaneous – Each codeword may be decoded immediately after its reception since it does not depend on any codewords to come.**

  => No codeword may be the 'starting' of any other codeword

- **'Bad' example:**
  - **Codewords: A - '0' ; B - '01' ; C - '11' ; D - '00' , E - '10'**
  - **Bitstream: 0000110 ...**
  - **Decoding: AAAACA ; DDCA ; ADBE ; ...**

Huffman (VLC) Coding

Símbolo	Probabilidade redução 0	Palav. Código redução 0	Probabilidade redução 1	Palav. Código redução 1
A	0.7	0	0.7	0
B	0.2	1 0	0.3	1
C	0.1	1 1		

Huffman coding allows obtaining a code with an average number of bits per symbol as close as desired to the source entropy.

But this requires knowledge on the source statistics, i.e., symbol probabilities.

Entropy = 1.157 bit/symbol
($H = \Sigma\ p_i \log_2 (1/p_i)$ bit/symbol)

Average code length = 1.3 bit/symbol

Efficiency = 1.157/1.3 = 89%

Huffman Coding: 2ª Order Extension

2nd extension	Source		Reduction 1		Reduction 2	
AA	0.49	1	0.49	1	0.49	1
AB	0.14	000	0.14	000	0.14	000
BA	0.14	001	0.14	001	0.14	001
AC	0.07	0100	0.07	0100	0.07	0100
CA	0.07	0101	0.07	0101	0.07	0101
BB	0.04	0111	0.04	0111	0.05	0110
BC	0.02	01101	0.03	01100	0.04	0111
CB	0.02	011000	0.02	01101		
CC	0.01	011001				

2nd extension	Red. 3		Red. 4		Red. 5		Red. 6		Red. 7	
AA	0.49	1	0.49	1	0.49	1	0.49	1	0.51	0
AB	0.14	000	0.14	000	0.23	01	0.28	00	0.49	1
BA	0.14	001	0.14	001	0.14	000	0.23	01		
AC	0.09	011	0.14	010	0.14	001				
CA	0.07	0100	0.09	011						
BB	0.07	0101								
BC										
CB										
CC										

Entropy = 1.157 bit/symbol

Average code length for 2nd order extension = 2.33 bit/extension symbol

Average code length = 2.33/2 = 1.165 bit/symbol

Efficiency = 1.157/1.165 = 99,3 %

Modified Huffman Method: Design Options

- **Black and White Coding Tables** - Due to their very different statistics, MHM uses separate Huffman coding tables for the black and white runs; with this solution, keeping the tone synchronism is essential for correct decoding.

- **Coding Long Runs** - To reduce the dimension of the Huffman tables, simplifying the implementations, runs longer than 63 pixels are coded in a different way. For these runs, their length is represented using 2 codewords: a *make-up code* multiple of 64 and a *terminating code* lower than 64.

$$\text{Run} = \text{Make-up Code} \times 64 + \text{Terminating Code}$$ (e.g. $739 = 11 \times 64 + 35$)

- The maximum value for the compression factor is set by the Information Theory as

$$CF_{max} = 1/H_{pixel} = (run_{white} + run_{black}) / (H_{white} + H_{black})$$

assuming that different codeword tables are used for black and white runs as their statistics are rather different.

MHM: Terminating Codes

White run length	*Code word*	*Black run length*	*Code word*
0	00110101	0	0000110111
1	000111	1	010
2	0111	2	11
3	1000	3	10
4	1011	4	011
5	1100	5	0011
6	1110	6	0010
7	1111	7	00011
8	10011	8	000101
9	10100	9	000100
10	00111	10	0000100
11	01000	11	0000101
12	001000	12	0000111
13	000011	13	00000100
14	110100	14	00000111
15	110101	15	000011000
16	101010	16	0000010111
17	101011	17	0000011000
18	0100111	18	0000001000
19	0001100	19	00001100111

... 63

MHM: Make-up Codes

White run lengths	*Code word*	*Black run lengths*	*Code word*
64	11011	64	0000001111
128	10010	128	000011001000
192	010111	192	000011001001
256	0110111	256	000001011011
320	00110110	320	000000110011
384	00110111	384	000000110100
448	01100100	448	000000110101
512	01100101	512	0000001101100
576	01101000	576	0000001101101
640	01100111	640	0000001001010
704	011001100	704	0000001001011
768	011001101	768	0000001001100
832	011010010	832	0000001001101
896	011010011	896	0000001110010
960	011010100	960	0000001110011
1024	011010101	1024	0000001110100
1088	011010110	1088	0000001110101
1152	011010111	1152	0000001110110
1216	011011000	1216	0000001110111
1280	011011001	1280	0000001010010
1344	011011010	1344	0000001010011
1408	011011011	1408	0000001010100
1472	010011000	1472	0000001010101
1536	010011001	1536	0000001011010
1600	010011010	1600	0000001011011
1664	011000	1664	0000001100100
1728	010011011	1728	0000001100101
EOL	000000000001	EOL	000000000001

ITU-T Fax Test Images

1

THE SLEREXE COMPANY LIMITED

SAPORS LANE · BOOLE · DORSET · BH 25 8 ER

TELEPHONE BOOLE (945 13) 51617 · TELEX 123456

Our Ref. 350/PJC/EAC 18th January, 1972.

Dr. P.N. Cundall,
Mining Surveys Ltd.,
Holroyd Road,
Reading,
Berks.

Dear Pete,

Permit me to introduce you to the facility of facsimile transmission.

In facsimile a photocell is caused to perform a raster scan over the subject copy. The variations of print density on the document cause the photocell to generate an analogous electrical video signal. This signal is used to modulate a carrier, which is transmitted to a remote destination over a radio or cable communications link.

At the remote terminal, demodulation reconstructs the video signal, which is used to modulate the density of print produced by a printing device. This device is scanning in a raster scan synchronised with that at the transmitting terminal. As a result, a facsimile copy of the subject document is produced.

Probably you have uses for this facility in your organisation.

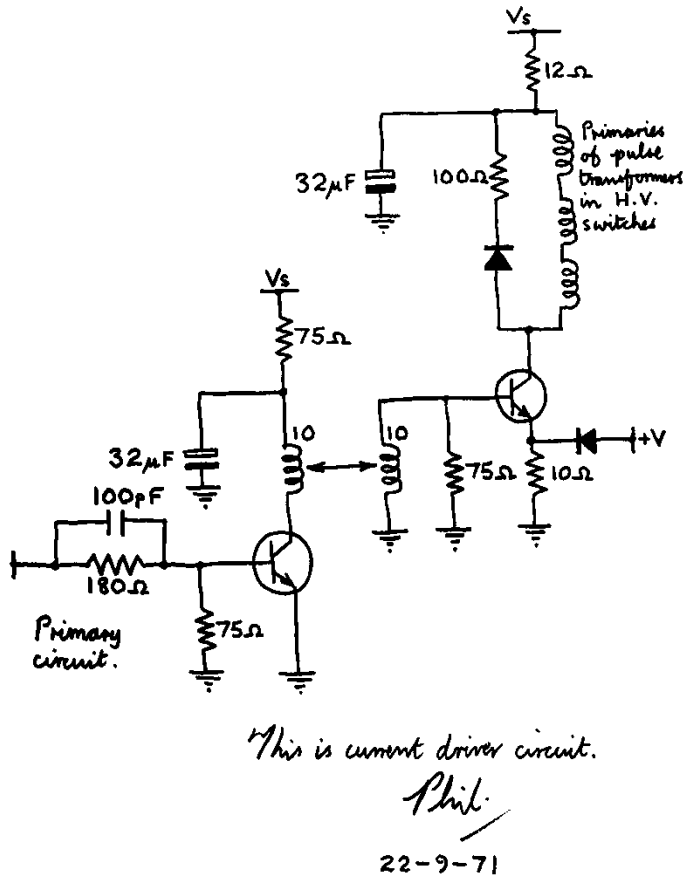
Yours sincerely,

Phil.

P.J. CROSS
Group Leader - Facsimile Research

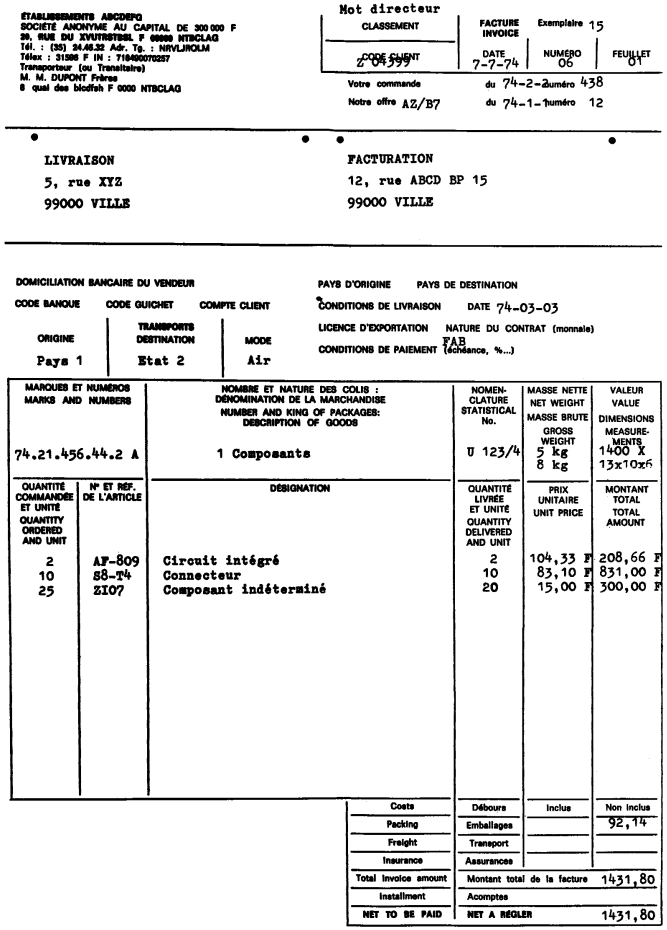
Registered in England: No. 2088
Registered Office: 60 Vicars Lane, Ilford, Essex.

2



ITU-T Fax Test Images

3



ÉTABLISSEMENTS ABCDEFG
SOCIÉTÉ ANONYME AU CAPITAL DE 300 000 F
20, RUE DU XVIITRIEME F 00000 NTBCLAG
Tél. : (35) 24.46.32 Adr. Tg. : NRVLJROLM
Télex : 31598 F IN : 716400070257
Transporteur (ou Transitaire)
M. M. DUPONT Frères
8 quai des biodfeh F 0000 NTBCLAG

Mot directeur

CLASSEMENT	FACTURE INVOICE	Exemplaire 15	
CODE CLIENT 2 05399	DATE 7-7-74	NUMERO 06	FEUILLET 01

Votre commande du 74-2-2 numéro 438
Notre offre AZ/B7 du 74-1-1 numéro 12

LIVRAISON
5, rue XYZ
99000 VILLE

FACTURATION
12, rue ABCD BP 15
99000 VILLE

DOMICILIATION BANCAIRE DU VENDEUR
CODE BANQUE | CODE GUICHET | COMPTE CLIENT

PAYS D'ORIGINE | PAYS DE DESTINATION
CONDITIONS DE LIVRAISON | DATE 74-03-03
LICENCE D'EXPORTATION | NATURE DU CONTRAT (monnaie) FAB
CONDITIONS DE PAIEMENT (échéance, %...)

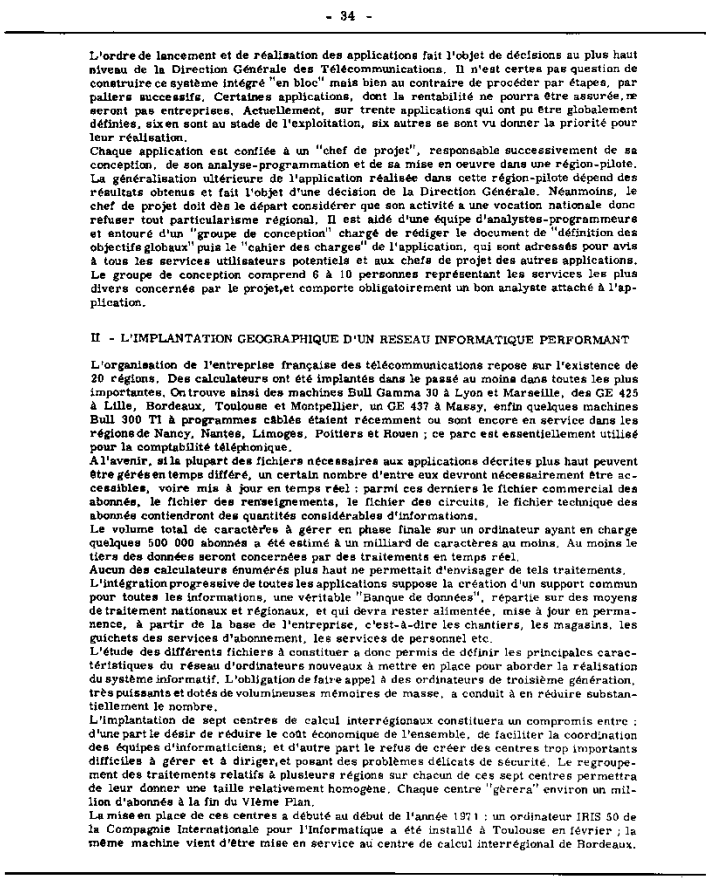
ORIGINE	TRANSPORTS DESTINATION	MODE
Pays 1	Etat 2	Air

MARQUES ET NUMEROS MARKS AND NUMBERS	NOMBRE ET NATURE DES COLIS : DENOMINATION DE LA MARCHANDISE NUMBER AND KING OF PACKAGES: DESCRIPTION OF GOODS	NOMENCLATURE STATISTICAL No.	MASSE NETTE NET WEIGHT / MASSE BRUTE GROSS WEIGHT	VALEUR VALUE / DIMENSIONS MEASUREMENTS
74.21.456.44.2 A	1 Composants	U 123/4	5 kg / 8 kg	1400 X 13x10x6

QUANTITÉ COMMANDÉE ET UNITÉ QUANTITY ORDERED AND UNIT	N° ET RÉF. DE L'ARTICLE	DÉSIGNATION	QUANTITÉ LIVRÉE ET UNITÉ QUANTITY DELIVERED AND UNIT	PRIX UNITAIRE UNIT PRICE	MONTANT TOTAL TOTAL AMOUNT
2	AF-809	Circuit intégré	2	104,33 F	208,66 F
10	88-T4	Connecteur	10	83,10 F	831,00 F
25	2107	Composant indéterminé	20	15,00 F	300,00 F

Costs	Débours	Inclus	Non inclus	
Packing	Emballages		92,14	
Freight	Transport			
Insurance	Assurances			
Total invoice amount	Montant total de la facture			1431,80
Installment	Acomptes			
NET TO BE PAID	NET A RÉGLER			1431,80

4



- 34 -

L'ordre de lancement et de réalisation des applications fait l'objet de décisions au plus haut niveau de la Direction Générale des Télécommunications. Il n'est certes pas question de construire ce système intégré "en bloc" mais bien au contraire de procéder par étapes, par paliers successifs. Certaines applications, dont la rentabilité ne pourra être assurée, ne seront pas entreprises. Actuellement, sur trente applications qui ont pu être globalement définies, six en sont au stade de l'exploitation, six autres se sont vu donner la priorité pour leur réalisation.
Chaque application est confiée à un "chef de projet", responsable successivement de sa conception, de son analyse-programmation et de sa mise en oeuvre dans une région-pilote. La généralisation ultérieure de l'application réalisée dans cette région-pilote dépend des résultats obtenus et fait l'objet d'une décision de la Direction Générale. Néanmoins, le chef de projet doit dès le départ considérer que son activité a une vocation nationale donc refuser tout particularisme régional. Il est aidé d'une équipe d'analystes-programmeurs et entouré d'un "groupe de conception" chargé de rédiger le document de "définition des objectifs globaux" puis le "cahier des charges" de l'application, qui sont adressés pour avis à tous les services utilisateurs potentiels et aux chefs de projet des autres applications. Le groupe de conception comprend 6 à 10 personnes représentant les services les plus divers concernés par le projet, et comporte obligatoirement un bon analyste attaché à l'application.

II - L'IMPLANTATION GEOGRAPHIQUE D'UN RESEAU INFORMATIQUE PERFORMANT

L'organisation de l'entreprise française des télécommunications repose sur l'existence de 20 régions. Des calculateurs ont été implantés dans le passé au moins dans toutes les plus importantes. On trouve ainsi des machines Bull Gamma 30 à Lyon et Marseille, des GE 425 à Lille, Bordeaux, Toulouse et Montpellier, un GE 437 à Massy, enfin quelques machines Bull 300 T1 à programmes câblés étaient récemment ou sont encore en service dans les régions de Nancy, Nantes, Limoges, Poitiers et Rouen ; ce parc est essentiellement utilisé pour la comptabilité téléphonique.
A l'avenir, si la plupart des fichiers nécessaires aux applications décrites plus haut peuvent être gérés en temps différé, un certain nombre d'entre eux devront nécessairement être accessibles, voire mis à jour en temps réel : parmi ces derniers le fichier commercial des abonnés, le fichier des renseignements, le fichier des circuits, le fichier technique des abonnés contiendront des quantités considérables d'informations.
Le volume total de caractères à gérer en phase finale sur un ordinateur ayant en charge quelques 500 000 abonnés a été estimé à un milliard de caractères au moins. Au moins le tiers des données seront concernées par des traitements en temps réel.
Aucun des calculateurs énumérés plus haut ne permettait d'envisager de tels traitements.
L'intégration progressive de toutes les applications suppose la création d'un support commun pour toutes les informations, une véritable "Banque de données", répartie sur des moyens de traitement nationaux et régionaux, et qui devra rester alimentée, mise à jour en permanence, à partir de la base de l'entreprise, c'est-à-dire les chantiers, les magasins, les guichets des services d'abonnement, les services de personnel etc.
L'étude des différents fichiers à constituer a donc permis de définir les principales caractéristiques du réseau d'ordinateurs nouveaux à mettre en place pour aborder la réalisation du système informatif. L'obligation de faire appel à des ordinateurs de troisième génération, très puissants et dotés de volumineuses mémoires de masse, a conduit à en réduire substantiellement le nombre.
L'implantation de sept centres de calcul interrégionaux constituera un compromis entre : d'une part le désir de réduire le coût économique de l'ensemble, de faciliter la coordination des équipes d'informaticiens; et d'autre part le refus de créer des centres trop importants difficiles à gérer et à diriger, et posant des problèmes délicats de sécurité. Le regroupement des traitements relatifs à plusieurs régions sur chacun de ces sept centres permettra de leur donner une taille relativement homogène. Chaque centre "gèrera" environ un million d'abonnés à la fin du VIème Plan.
La mise en place de ces centres a débuté au début de l'année 1971 : un ordinateur IRIS 50 de la Compagnie Internationale pour l'Informatique a été installé à Toulouse en février ; la même machine vient d'être mise en service au centre de calcul interrégional de Bordeaux.

Photo n° 1 - Document très dense lettre 1,5mm de haut -
Restitution photo n° 9

ITU-T Fax Test Images

5

Cela est d'autant plus valable que $T\Delta f$ est plus grand. A cet égard la figure 2 représente la vraie courbe donnant $|\phi(f)|$ en fonction de f pour les valeurs numériques indiquées page précédente.

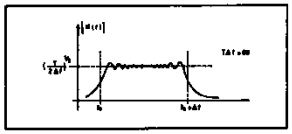


FIG. 2

Dans ce cas, le filtre adapté pourra être constitué, conformément à la figure 3, par la cascade :

— d'un filtre passe-bande de transfert unité pour $f_0 \leqslant f \leqslant f_0+\Delta f$ et de transfert quasi nul pour $f < f_0$ et $f > f_0+\Delta f$, filtre ne modifiant pas la phase des composants le traversant ;

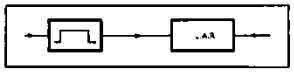


FIG. 3

— filtre suivi d'une ligne à retard (LAR) dispersive ayant un temps de propagation de groupe T_R décroissant linéairement avec la fréquence f suivant l'expression :

$$T_R = T_0+(f_0-f)\frac{T}{\Delta f} \quad (\text{avec } T_0 > T)$$

(voir fig. 4).

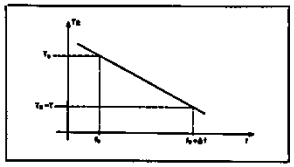


FIG. 4

telle ligne à retard est donnée par :

$$\varphi = -2\pi\int_0^f T_R\,df$$

$$\varphi = -2\pi\left[T_0+\frac{f_0 T}{\Delta f}\right]f+\pi\frac{T}{\Delta f}f^2$$

Et cette phase est bien l'opposé de $/\phi(f)$, à un déphasage constant près (sans importance) et à un retard T_0 près (inévitable).

Un signal utile $S(t)$ traversant un tel filtre adapté donne à la sortie (à un retard T_0 près et à un déphasage près de la porteuse) un signal dont la transformée de Fourier est réelle, constante entre f_0 et $f_0+\Delta f$, et nulle de part et d'autre de f_0 et de $f_0+\Delta f$, c'est-à-dire un signal de fréquence porteuse $f_0+\Delta f/2$ et dont l'enveloppe a la forme indiquée à la figure 5, où l'on a représenté simultanément le signal $S(t)$ et le signal $S_1(t)$ correspondant obtenu à la sortie du filtre adapté. On comprend le nom de récepteur à compression d'impulsion donné à ce genre de filtre adapté : la « largeur » (à 3 dB) du signal comprimé étant égale à $1/\Delta f$, le rapport de compression est de $\frac{T}{1/\Delta f} = T\Delta f$

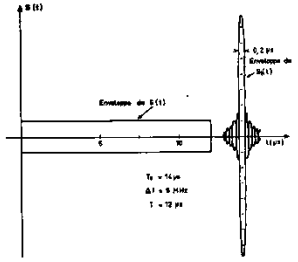
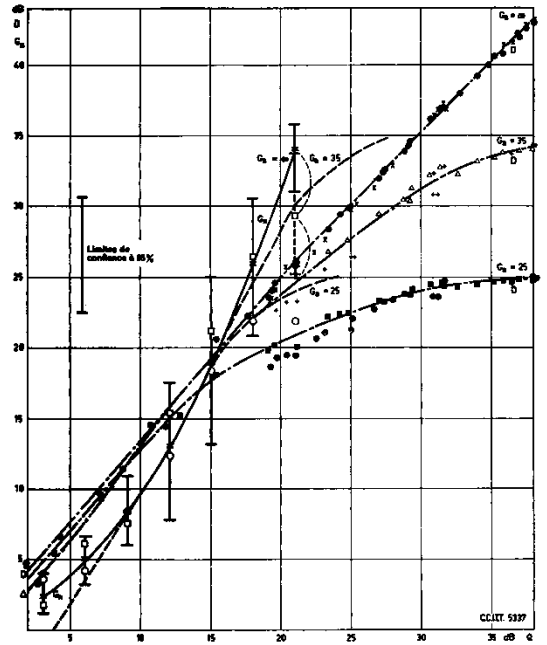


FIG. 5

On saisit physiquement le phénomène de compression en réalisant que lorsque le signal $S(t)$ entre dans la ligne à retard (LAR) la fréquence qui entre la première à l'instant 0 est la fréquence basse f_0, qui met un temps T_0 pour traverser. La fréquence f entre à l'instant $t = (f-f_0)\frac{T}{\Delta f}$ et elle met un temps $T_0-(f-f_0)\frac{T}{\Delta f}$ pour traverser, ce qui la fait ressortir à l'instant T, également. Ainsi donc, le signal $S(t)$

6



Courbes adaptées G_a (essais subjectifs) pour
$G_a = \infty$ $G_a = 35$dB $G_a = 25$dB
[10] ——— × □ ○
[23] — — — —

Points calculés $D\,(Q, G_a)$ pour
$G_a = \infty$ $G_a = 25$ $G_a = 35$ dB
● ■ ▲ – dans la partie montante
× ● + – dans la partie descendante

Courbes ——— — — — — $D\,(Q, G_a)$

FIGURE 3

ITU-T Fax Test Images

7

ＣＣＩＴＴの概要

沿革

ＣＣＩＴＴは、「国際電気通信諮問委員会（ＩＴＵ）」の四つの常設機関（事務総局、国際周波数登録委員会、ＣＣＩＲ、ＣＣＩＴＴ）の一つとして、世界の国際通信上の諸問題を真先に取上げ、その解決方法を見出して行く重要な機関である。日本名は、「国際電信電話諮問委員会」と称する。

ＣＣＩＴＴの前身は、ＣＣＩＦ（国際電話諮問委員会）とＣＣＩＴ（国際電信諮問委員会）である。ＣＣＩＦは、１９２４年にヨーロッパに「国際長距離電話通信諮問委員会」が設置され、これが１９２５年のパリ電信電話会議のとき、正式に、「国際電話諮問委員会」として万国電信連合の公式機関となったものである。ＣＣＩＴは、同じく１９２５年の会議のとき、ＣＣＩＦと併立するものとして設置された。

そして、ＣＣＩＦは、１９５８年の12月に第８回総会が開催されたのち、ＣＣＩＴは、同年同月に第８回総会が開催されたのち、併合されて現在のＣＣＩＴＴとなった。このＣＣＩＴＴは、ＣＣＩＦとＣＣＩＴが解散した直後、第１回総会を開催し、第２回総会は、１９６０年にニューデリーで、第３回総会は、１９６４年、ジュネーブで、第４回総会は、１９６８年、アルゼンチンで開催された。

ＣＣＩＦとＣＣＩＴが合併したのは、有線電気通信の分野、とくに伝送路について電信回線と電話回線とを技術的に分ける意味がなくなってきたこと、各国とも大体において、電信部門と電話部門は同一組織内にあること、ＣＣＩＦの事務局とＣＣＩＴＴの事務局の合併による能率増進等がおもな理由であった。

ＣＣＩＴＴは、上述のように、ヨーロッパ内の国ぐにによって、ヨーロッパ内の電信・電話の技術・運用・料金の基準を定め、あるいは統一をはかってきたので、現在でも、その影響を受け、会各参加国は、ヨーロッパの国が多く、ヨーロッパで生起する問題の研究が多い。たとえば、１９６０年のＣＣＩＴＴ勧告の中で、技術上配慮する距離は約2,500kmであったが、これはヨーロッパ内領域を想定したものである。

しかしながら、１９５６年９月に敷設された大西洋横断電話ケーブルは、大陸間電話通信の自動化および半自動化への技術的可能性を与え、ＣＣＩＴＴがこの問題を取り上げるに及び、ＣＣＩＴＴの性格は飛躍的に増びるに至った。この汎世界的性格は第２次世界大戦後目ざましくなったアジア・アフリカ植民地の独立に伴ってＩＴＵの構成員の中に これらの国が加わり、ＩＴＵの中に新しい意見が導入されたことにも起因して、技術面、政治面の双方から導入されてきた。ＣＣＩＴＴの汎世界化は、１９６０年の第２回総会がニューデリーで開催されたことにもあらわれている。この総会までは、ＣＣＩＴ、ＣＣＩＦのいずれにしろ、アメリカやアジアで総会が開催されたことがなく、ＣＣＩＴＴ委員長も、ニューデリー総会の準備文書で、この点には注目すべきであるとのべている。

任務

ＩＴＵは、全権委員会議、主管庁会議を始めとして、七つの機関をもち、それぞれの機関の権限と任務は国際電気通信条約に明記されている。そこで条約を参照してみるならば、ＣＣＩＴＴの任務は、つぎのとおりとなっている。

「国際電信電話諮問委員会（ＣＣＩＴＴ）は、電信および電話に関する技術、運用および料金の問題について研究し、および意見を表明することを任務とする。」（１９６５年モントルー条約第１９７号）

「各国際諮問委員会は、その任務の遂行に当たって、新しい国または発展の途上にある国における地域的および国際的分野にわたる電気通信の創設、発達および改善に直接関連のある問題について研究し、および意見を作成するように妥当な注意を払わなければならない。」（同第１８８号）

「各国際諮問委員会は、また、関係国の要請に基づき、その国内電気通信の問題について研究し、かつ、勧告を行なうことができる。」（同第１８９号）

上記第１８７号と第１８８号にいわれる「意見」とは、フランス語の Avis から訳したもので、英語では「勧告（Recommendation）」となっている。ＣＣＩＴＴの表示する意見は、国際法的には強制力をもたないものであって、この点が、条約、電信規則、電話規則等各国を拘束する力をもっているものと異なる。もっとも意見とは称しても、技術的分野では、電信規則のごとき、各国政府が承認してその内容を実施する強制規則をもたないので、実際にある機器の仕様を定める場合には、多くの国の意見が統一されたこの「意見」に従わなければ、円滑な国際通信を行なうことができない場合が多い。この意見（または勧告）は、国際通信を行なう場合各国が直面する問題について、具体的意見を表明するもので、たとえば、大陸間ケーブルや大陸間通話を半自動化しようとする場合、その信号方式や取り扱う通話の種類および料金は、どのようにするかを研究して意見を表明する。したがって、ＣＣＩＴＴの活動は、つねに時代の最先端を行くもので、ＣＣＩＴＴの活動方向は、そのまま世界の国際通信の活動方向であるともいえる。

この意見は、また、電信規則以下のその他の規則のごとく、数年以上の間隔をもって開催される主管庁会議というような大会議の決定をまたなくても変更することができ、また、その改正も容易であるので、現在のように進歩の早い国際通信界では、関係国の意見を統一した国際的見解としては非常に便利である。

8

memorandum

TO: G.V. Smith, Project Planning
FROM: A.P. Spriggs, Research
DATE: 1-9-71
SUBJECT: 2041

We know that, where possible, data is reduced to alphanumeric form for transmission by communication systems. However, this can be expensive, and also some data must remain in graphic form. For example, we cannot key-punch an engineering drawing or weather map.

I think we should realise that high speed facsimile transmissions are needed to overcome our problems in efficient graphic data communication. We need research into graphics data compression.

Any comments?

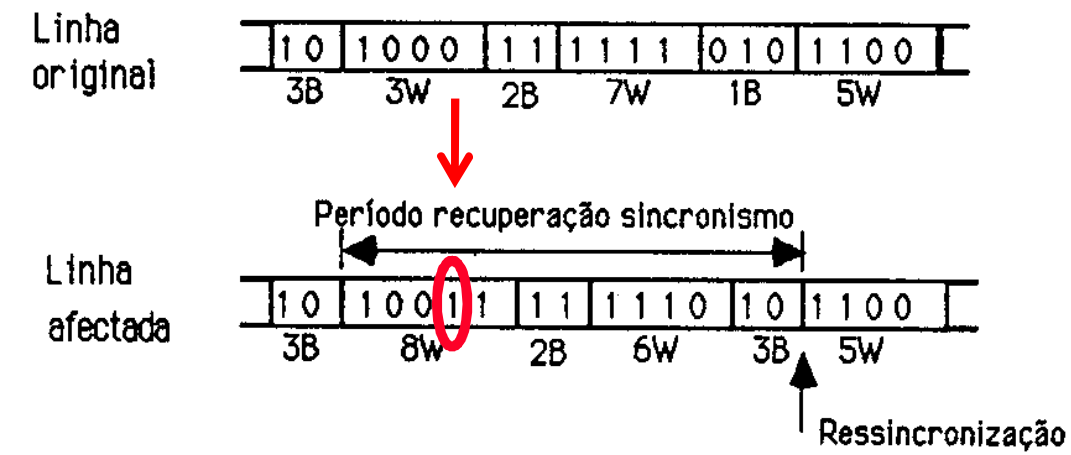
P Elbert.

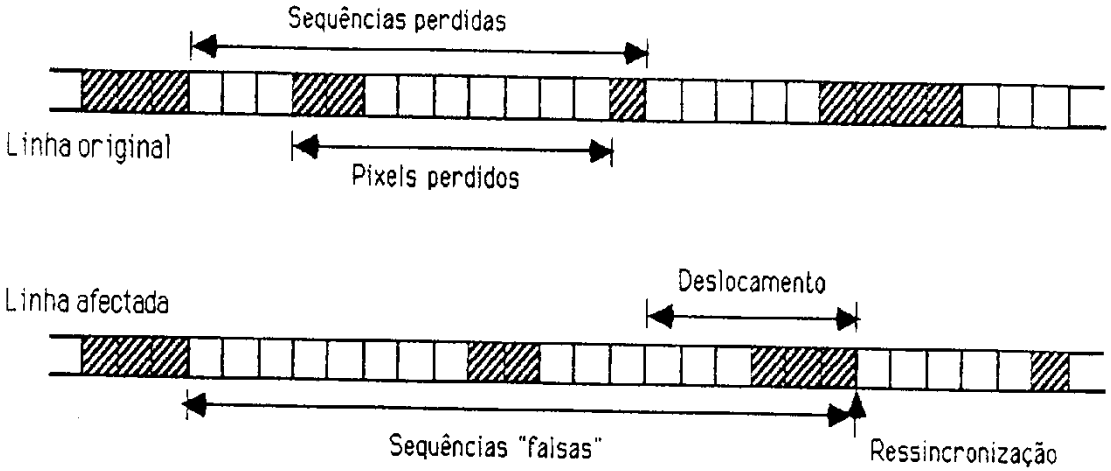
WELL, WE ASKED FOR IT !

MHM: Compression Factor

Compression Efficiency for the ITU-T Fax Test Images					
Doc.	Avg. white run	Avg. black run	Entropy for white runs	Entropy for black runs	CF_{max}
1	156.3	6.793	5.451	3.592	18.02
2	257.1	14.31	8.163	4.513	21.41
3	89.81	8.515	5.688	3.572	10.62
4	39.00	5.674	4.698	3.124	5.712
5	79.16	6.986	5.740	3.328	9.5
6	138.5	8.038	6.204	3.641	14.89
7	45.32	4.442	5.894	3.068	5.553
8	85.68	70.87	6.862	5.761	12.4

Compression Efficiency for the ITU-T Fax Tests Images using MHM						
Doc.	Avg. white runs	Avg. black runs	Entropy for white runs	Entropy for black runs	CF_{max}	CF_{real}
1	134.6	6.79	5.23	3.592	16.02	15.16
2	167.9	14.02	5.989	4.457	17.41	16.67
3	71.5	8.468	5.189	3.587	9.112	8.35
4	36.38	5.673	4.574	3.126	5.461	4.911
5	66.41	6.966	5.280	3.339	8.513	7.927
6	90.65	8.001	5.063	3.651	11.32	10.78
7	39.07	4.442	5.320	3.068	5.188	4.99
8	64.30	60.56	4.427	5.31	11.52	8.665





MHM: Resilience to Errors

The period to recover the synchronism is defined as the number of bits between the starting of the corrupted codeword and the end of the codeword where the synchronism is recovered.

Digital Coding of Bilevel Images

GROUP 3 FAX

- **MODIFIED HUFFMAN METHOD (MHM) – Unidimensional coding method based on the coding of the lenght of alternate black and white pixel runs using Huffman coding.**

GROUP 4 FAX (also Group 3 options)

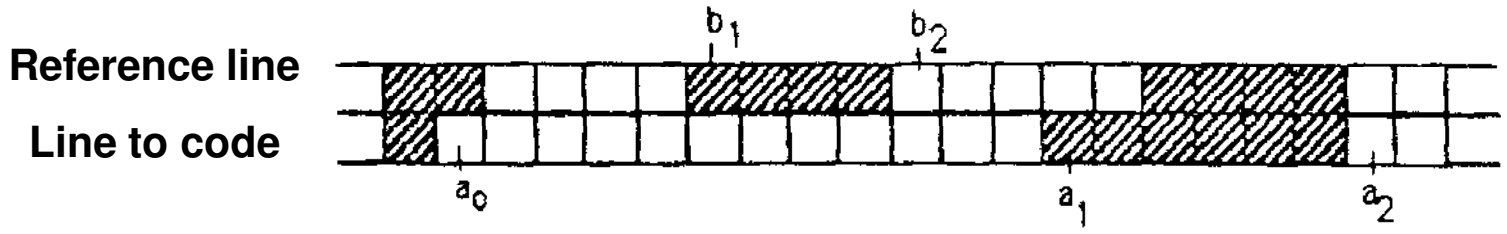
- **MODIFIED READ METHOD (MRM) – Bidimensional coding method based on the coding of the variations of the positions of tone transition pixels (black-white or white-black) in relation to the previous line; unidimensional coding may be used every k lines.**
- **MODIFIED-MODIFIED READ METHOD (MMRM) – Similar to MRM but without periodic unidimensional coding.**

Modified Read Method: the Symbols

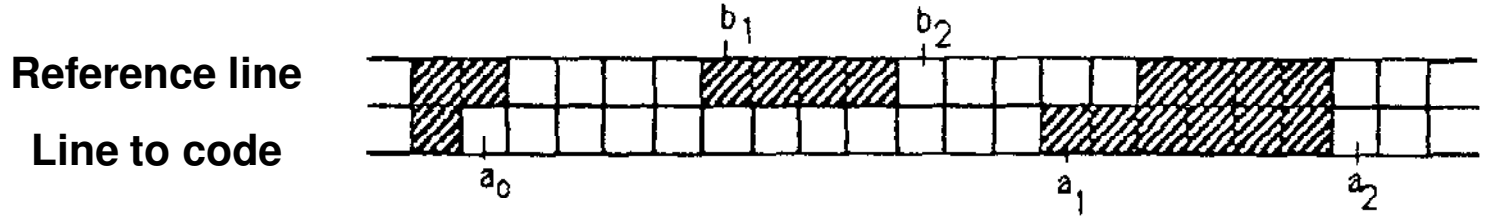
The Modified READ (*relative addressing*) Method (MRM) exploits the vertical redundancy in the image (in addition to the horizontal redundancy) to achieve higher compression factors.

MRM is a line by line coding method where the position of each variation element in the line (a_0, a_1, a_2, b_1, b_2) to code is coded:

- **Using as reference the position of the corresponding variation element in the reference (previous) line (vertical redundancy) or**
- **Using as reference the previous variation element in the line to code (horizontal redundancy)**



MRM: Variation Elements

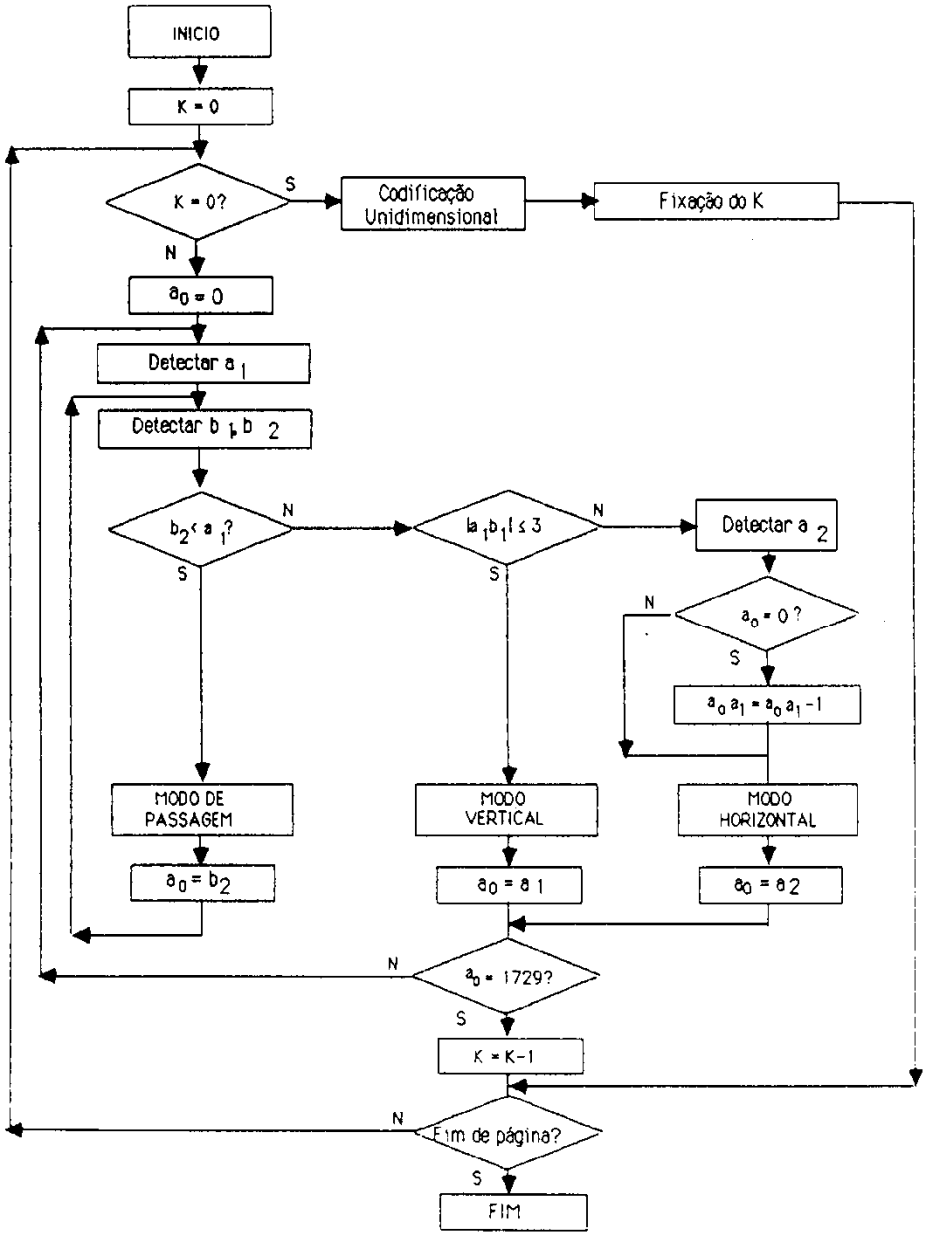


- **A variation element is a pixel which tone is different from the tone of the previous variation element in the same line.**
- **The MRM algorithm uses 5 variation elements located in the line to code as well as in the reference (previous) line:**
  - a_0 **– reference or starting element in the line to code; its position is defined by the preceeding coding mode. At the starting of the line to code, a_0 is located in a virtual white variation element placed immediately before the first pixel of the line to code**
  - a_1 **– variation element immediately after a_0 in the line to code; this element has a tone opposite to a_0 and it is the next variation element to code**
  - a_2 **– first variation element at the right of a_1**
  - b_1 **– first variation element in reference line at the right of a_0 with the same tone of a_1**
  - b_2 **– first variation element at the right of b_1**

MRM: Coding Modes

- **VERTICAL MODE – Used when there is a good correlation between the reference line and the line to code – the position of a_1 is coded relative to the position of b_1. The distance a_1-b_1 may take 7 values: 0, ± 1, ± 2 e ± 3.**
- **PASS MODE – Serves to skip a black run in the reference line – this mode happens when the position of b_2 is at the left of the position a_1; only one codeword is needed.**
- **HORIZONTAL MODE – Used when there is a black run in the line to code without sufficient correlation with the reference line – used when the vertical mode cannot be used; the distances a_0-a_1 and a_1-a_2 are sent.**
- **WITHOUT COMPRESSION MODE – Uses the PCM values (1 sample, 1 bit) allowing that, for very detailed zones, the number of code bits is never higher than the number of samples and, thus, PCM bits.**

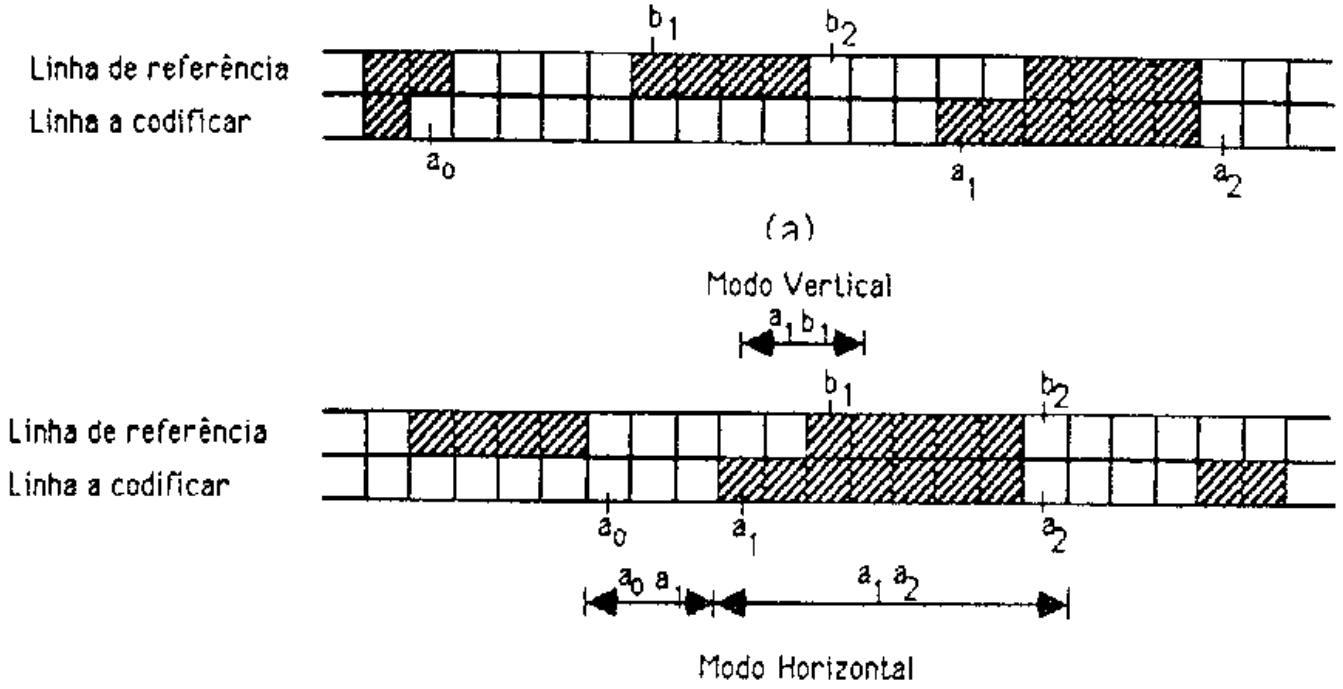
EXTENSAO 2D



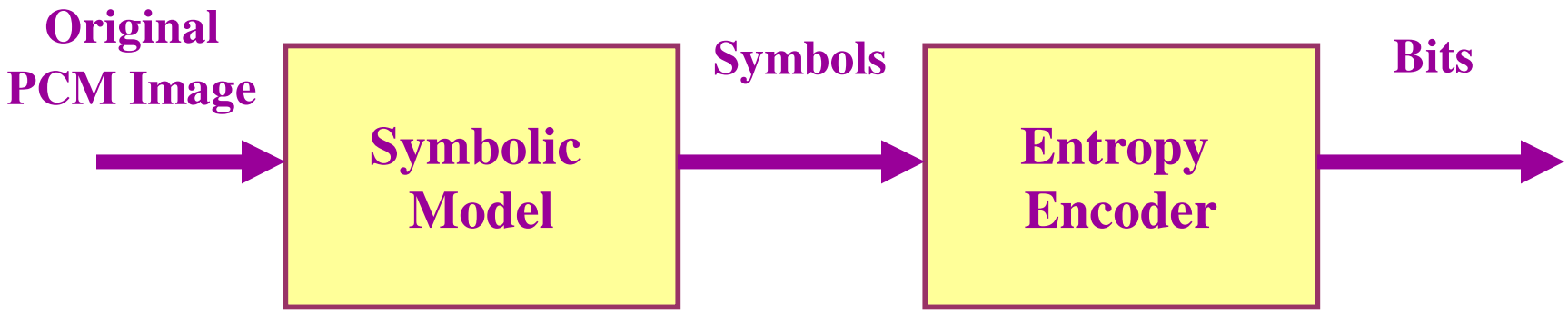
MRM Coding Process

Modified READ Method: Stopping Error Propagation ...

To minimize the vertical propagation of damages caused by transmission errors, no more than *k-1* successive lines are coded using the bidimensional procedure. This means that each *kth* line is coded using the MHM unidimensional procedure.



MRM Facsimile Coding



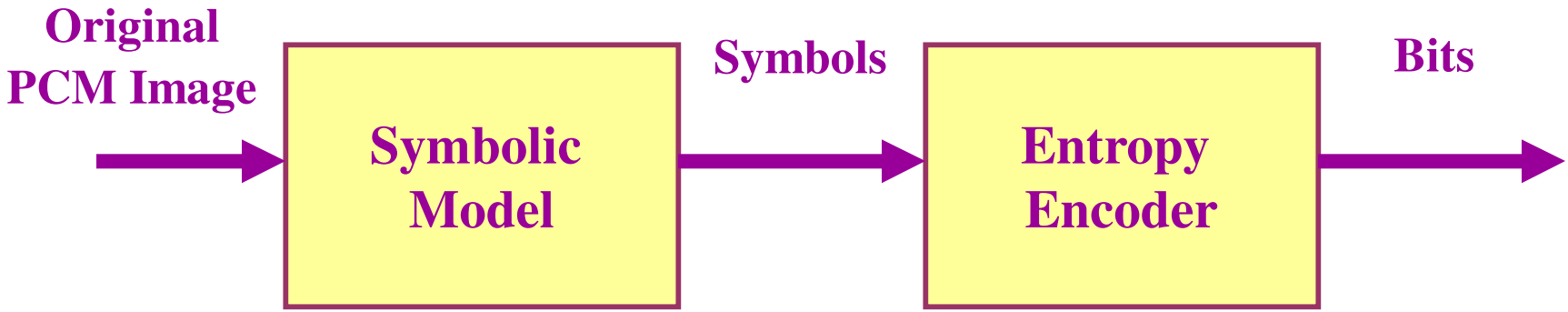
MRM Model: A facsimile image is represented as a sequence of depending lines, each of them represented as a sequence of symbols representing the BW and WB edges, using as references the edges in the previous or same line; periodically, one line is coded without exploiting the vertical redundancy, this means using the MHM model.

MHM and MRM: Comparing Performances

Number of bit/image						
	Low resolution (MSLT= 0 ms)			High resolution (MSLT= 0 ms)		
Doc.	MHM	MRM (k=2)	Gain %	MHM	MRM (k=4)	Gain %
1	149834	130684	12.8	299311	207660	30.6
2	137252	106851	22.1	274858	157163	42.8
3	260247	207584	20.2	520196	326297	37.3
4	432219	408261	5.5	864524	654436	24.3
5	273164	226285	17.2	546460	353172	35.4
6	204516	150572	26.4	409290	225879	44.8
7	426053	402333	5.6	851286	651643	23.5
8	251171	210457	16.2	502331	264029	47.4
Average	266807	227117	15.8	533532	355034	40

- MRM allows to achieve higher compression and thus lower transmission times; the reduction is larger for high resolution (7.7 pel/mm versus 3.85pel/mm) and may achieve almost 50 % for MSLT= 0 ms (MSLT is the *Minimum Scan Line Time*).
- MRM compression efficiency advantages are higher for less dense/detailed images (where there is more vertical redundancy to exploit).
- MRM is more complex than MHM.
- MRM is more sensitive to transmission errors.

MMRM Facsimile Coding



MMRM Model: A facsimile image is represented as a sequence of depending lines, each of them represented as a sequence of symbols representing the BW and WB edges, using as references the edges in the previous or same line (no periodic MHM coded line is inserted and also no EOLs are inserted).

Digital Coding of Bilevel Images: the k Factor

GROUP 3 FAX

k=1

- **MODIFIED HUFFMAN METHOD (MHM) – Unidimensional coding method based on the coding of the lenght of alternate black and white pixel runs using Huffman coding.**

GROUP 4 FAX (also Group 3 options)

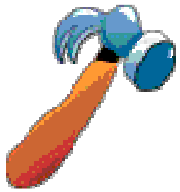
1 < k <∞

- **MODIFIED READ METHOD (MRM) – Bidimensional coding method based on the coding of the variations of the positions of tone transition pixels (black-white or white-black) in relation to the previous line; unidimensional coding may be used every *k* lines.**

k=∞

- **MODIFIED-MODIFIED READ METHOD (MMRM) – Similar to MRM but without periodic unidimensional coding.**

Transmission Errors



- Any transmission using the telephone lines must consider the effects of transmission errors.
- Typically, the more efficient are the coding methods, the more sensitive they are since every bit carries more information (on average). However, more efficient coding methods (achieving lower bitrates) suffer less transmission errors, leading to te so-called statistical protection effect.
- The receiver may detect the ocorrence of transmission errors and process the received signal to minimize the subjective effects in the decoded image of the errors.
- Errors may be detected when:
  - **Semantic condition:** The decoded line does not have the correct number of pixels, e.g. 1728 pixels/line for low resolution (MHM and MRM).
  - **Syntactic condition:** None of the codewords in the tables corresponds to the received sequence of bits (MHM e MRM).
  - **Syntactic condition:** The line to decode refers a run that does not exist in the reference line (MRM).

Minimizing the Subjective Impact of Errors: Error Concealment (1)

Error concealment corresponds to the process where the receiver creates data for the parts received in error (and for which no correction capability was available) while maximizing the subjective quality.

Error concealment is more important for MRM coding due to the vertical (and not only horizontal) propagation of errors.

Example error concealment techniques (with increasing complexity):

- **PRINT WHITE – The first erroneous line is printed white and all subsequent lines are printed white until a one-dimensional coded (MHM) line is correctly received.**
- **PRINT PREVIOUS LINE – The first erroneous line is replaced by the previous correctly received line and all subsequent lines are replaced by that line until a one-dimensional coded (MHM) line is correctly received.**

Minimizing the Subjective Impact of Errors: Error Concealment (2)

Error concealment techniques (by increasing complexity):

- **PRINT PREVIOUS LINE AND AFTER WHITE – The first erroneous line is replaced by the previous correctly received line and all subsequent lines are printed white until a one-dimensional coded (MHM) line is correctly received.**

- **NORMAL DECODE/PREVIOUS LINE – The first erroneous line is decoded and printed in the normal manner up to the point in the line where the error is detected. From this point, the remainder of the first erroneous line is replaced by the corresponding pixels in the previous line. The resultant line is then used as a new reference line and the process is repeated until a unidimensional coded (MHM) line is correctly decoded.**

Error Sensitivity Factor

MRM Error Sensitivity Factor (Doc. 1, 4 and 5)					
Resolution	Factor K	Method 1	Method 2	Method 3	Method 4
Normal	2	36.24	24.64	29.60	23.20
	3	34.03	40.89	31.01	27.76
High	4	66.55	49.23	55.16	54.49
	6	88.51	64.46	76.55	75.74
Average		56.33	44.80	48.08	45.32

ESF tends to decrease … Readability tends to increase ...

The Error Sensitivity Factor (ESF) corresponds to the average number of incorrect pixels for each transmission error.

Group 3 Fax Error Control

Group 3 fax basic configuration does not foresee the use of any error control techniques. But there are some extensions/tools …

However:

- **For MRM, the periodic transmission of unidimensionally coded lines targets the limitation of error propagation.**
- **Some faxs may ask for the retransmission of the page, if more than X lines are detected as erroneous.**
- **The initial protocol defines the transmission rate depending on the line conditions to limit the error rate.**

The Beauty or the Monster ?

- **A long hibernation – The deployment of fax has stressed the importance of standardization and influenced the way standardization is made nowadays.**
- **Democratization – The easiness to install and use a fax and its price have made it a very largely used equipment also for protest and revolutionary purposes, e.g. Tian amen up rise.**
- **Transparency – Its autonomy and initial transparency led to some problems and the consequent adoption of privacy protection technology, e.g. passwords, cryptography.**
- **The 'intruse' – Its widespread usage transformed it in a powerful and simple advertizing mechanism 'by force'. Technology and law responded ...**
- **Impunity ? – A communication system where there is no face and no voice may serve less proper purposes ...**

Bibliography

- **FAX - Digital Facsimile Technology & Applications,**
  K.McConnel, D.Bodson, R.Schaphorst, Artech House, 1992