

A entropia da informação como discriminante*

F. CARVALHO RODRIGUES

Abstract

We consider the possibility of using Shannon's quantity of information, also called entropy of information, to characterize defects in textiles and leathers. We apply the same concept to the measurement of randomness of fibers in paper and the automatic classification and counting of agricultural species.

We chose those economically important sectors to illustrate the scope of the method of measurement of information's entropy using standard vision equipment. It is demonstrated that the entropy of information affords the possibility of a single parameter judgment on the quality of products in a wide range of highly complex industrial operations.

INTRODUÇÃO

Entropia da informação passou desde a sua criação por Shannon com a cumplicidade de Von Neumann a ter um conjunto de significados que tornaram o conceito fugidio. Este facto tem como consequência um alheamento continuado em relação à possibilidade de medir a portanto utilizar um conceito basilar na teoria da informação.

Neste trabalho abordamos o problema da medição expedita, eficaz, precisa a rigorosa da entropia de um sistema quando a informação que emite é suportada pelo espectro electromagnético, designadamente, pela luz.

* LNETI, Laboratório Nacional de Engenharia a Tecnologia Industrial. IST, Instituto Superior Técnico.

Assim, após uma revisão breve do conceito de medição da entropia utilizando sistemas de visão convencionais a como ilustração da capacidade do método de medida é descrita a aplicação em três grandes sectores industriais a um das indústrias dos serviços: têxtil, curtumes, papel a em cartografia. Concluiu-se da eficácia do método de medida a da sua capacidade para determinação de zonas com variações relativamente ao comportamento desejado do processo industrial revelando assim defeitos nos produtos, nos processos de fabrico a na qualidade das matérias primas. Concluiu-se, ainda, que utilizando a variável informação e uma medida, a entropia da informação, esta pode ser utilizada como parâmetro discriminante em processos de enorTne complexidade, como o são os do fabrico de produtos das indústrias anteriormente mencionadas.

1. ENTROPIA DA INFORMAÇÃO. CONCEITO DE SHANNON

Sobre a multiplicidade e a variedade do que tem sido a definição da variável informação existe já literatura em que se enumeram, se registam notarialmente a se comparam os diversos exercícios que se destinaram a cristalizar o conceito que dá pelo nome de informação.

A questão, neste trabalho, fica enquadrada pela definição de Shannon ou se se quiser pela utilização que Shannon faz para conseguir uma quantidade de informação calculada a partir de um conjunto de mensagens. Para Shannon a quantidade de informação não fica agarrada às mensagens e a todas as ligações entre receptor e emissor no plano da semântica a do relativismo, antes fica definida pelas probabilidades que essas mensagens têm ou não de ser emitidas. Se se quiser trata-se de informação despojada de conhecimento.

E neste âmbito que a medida da quantidade de informação de **Shannon traduz uma propriedade intrínseca dos** sistemas que se traduz na célebre equação, da entropia da informação, H:

$$H = - \sum p_i \log p_i \quad (1)$$

em que p_i é a probabilidade da mensagem i . O suporte da mensagem é irrelevante para a definição.

O valor de H é característico do sistema num determinado estado macroscópico. A alteração quer da forma (na variante da informação) quer do estado produz modificação na medida que porventura se faça de H . Neste sentido a entropia da informação, como caracterizadora de um sistema pode ser utilizada como discriminante. Aliás, do mesmo modo que na variável espaço-tempo, o conceito de acontecimento num determinado instante divide o passado e o futuro, a na variável espaço podemos aferir a variação nas dimensões do sistema a na energia alterações do seu comportamento, a utilização da informação pode ser empregue para aquilatar a variabilidade de qualquer sistema. Para o fazer basta que tal como para as medidas do tempo, espaço a energia exista um método que leva à medida da quantidade de informação sinónimo de entropia de informação tal como é definida na expressão (1).

2. MEDIDA DA ENTROPIA DA INFORMAÇÃO

Para medir a entropia da informação de um sistema, de acordo com a expressão (1), basta medir a probabilidade de cada uma das mensagens que ele emite. Então para a medida, a natureza do suporte das mensagens é de grande importância, uma vez que a detecção da mensagem e a medida da sua probabilidade são os factores essenciais.

Nestas circunstâncias a medida da entropia da informação seja de que sistema for será sempre dentro de uma gama de mensagens da mesma natureza uma vez que não haverá sensores que permitam instantaneamente reconhecer as mensagens suportadas por todo o tipo de emissão corpuscular ou radiativa a avaliar a probabilidade de ocorrência de cada uma delas. Assim sendo a medida da entropia da informação só poderá ser utilizada para efectuar comparações quando diz respeito ao mesmo elenco de mensagens.

Uma vez assegurada a uniformidade atrás referida, a medida da entropia da informação, fica muito facilitada. Basta que exista um sensor que atribua uma probabilidade para cada mensagem.

3. MEDIDA DA ENTROPIA DA INFORMAÇÃO NO ESPECTRO VISÍVEL

Quando as mensagens que são escrutinadas têm por veículo a luz, a tarefa da medida da quantidade de informação fica bastante simplificada.

Uma câmara CCD, ligada a um sistema de digitalização de imagem, figura 1, permite medir a probabilidade da quantidade

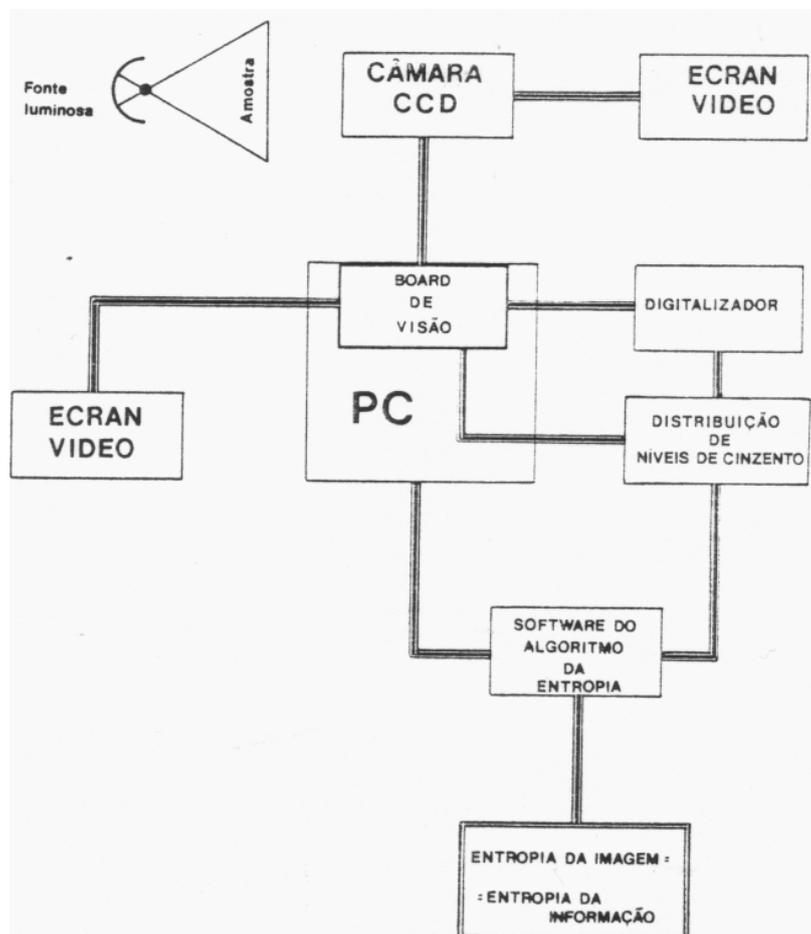


Figura 1

Diagrama de blocos para a medida da entropia de informação.

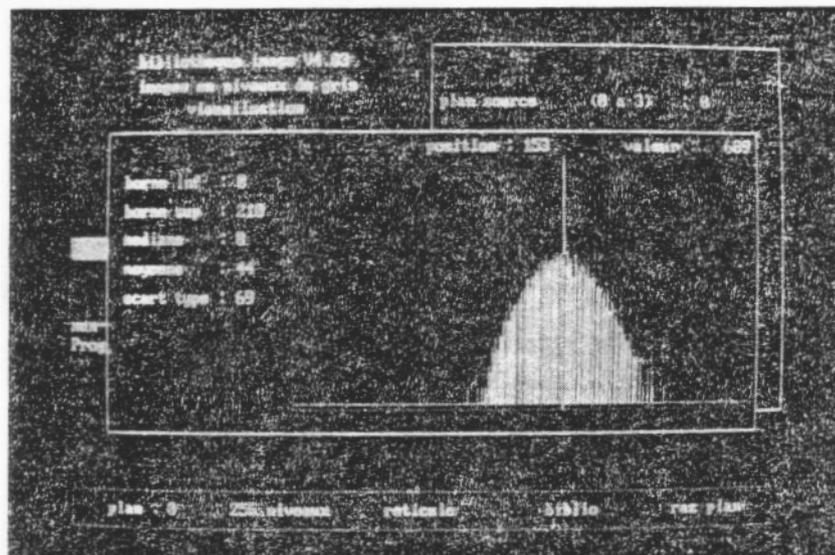


Figura 2

Distribuição típica de níveis de cinzento de uma imagem. Entropia: 2.243.

de luz emitida num espectro largo por uma área definida pela ampliação do sistema óptico a pelo pixel definido pela resolução do CCD. A quantidade luz na imagem, ou seja, o nível de cinzento de cada pixel é registado a uma função de distribuição de níveis de cinzento, figura 2, é obtida de imediato.

A partir da função de distribuição das mensagens, neste método a iluminação da imagem do sistema, o emprego da expressão de definição de entropia da informação permite, de imediato, o seu cálculo.

4. APLICAÇÕES DA ENTROPIA DA INFORMAÇÃO

A medida da entropia da informação de um sistema a partir da medida da mesma quantidade referida a uma sua imagem permite seguir aquele parâmetro, em tempo real, mesmo para sistemas com estruturas altamente complexas. Quer pela complexidade, quer

pela importância económica dos sectores industriais, os processos têxtil, do papel e dos curtumes são exemplos da aplicação ilustrativa da aplicabilidade da entropia de informação como discriminante na detecção de anomalias nos processos de fabrico e de defeitos no produto final.

4.1. Sector Têxtil

4.1.1. Penteação

Processos de fabrico, como é o caso da penteação da lã, passaram a poder ser completamente aferidos. Neste caso a interpretação, sempre ligada à entropia da informação, da ordem e da desordem que atribui à desordem um valor mais elevado da medida

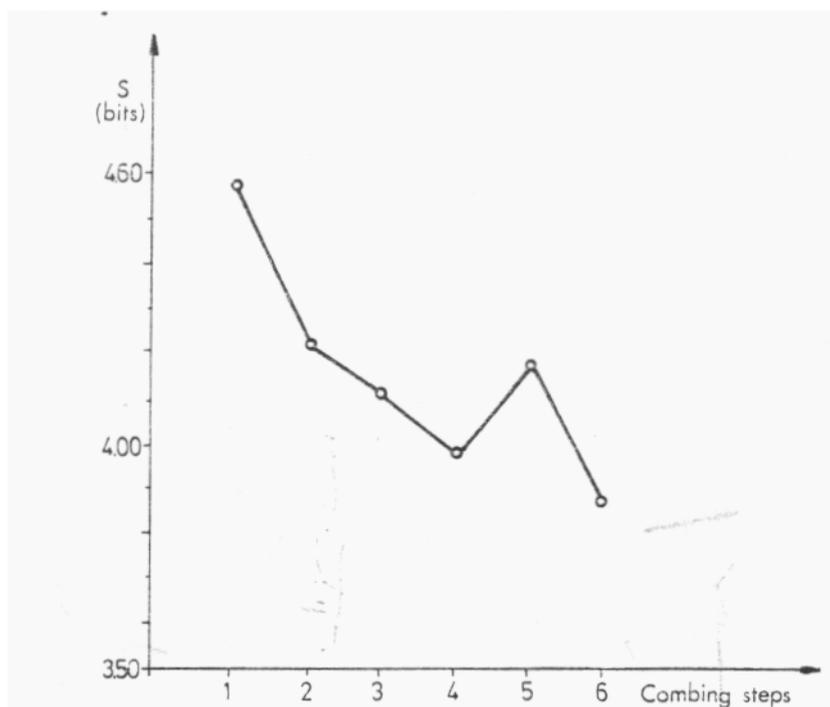


Figura 3

Evolução da entropia de informação ao longo da penteação da lã.

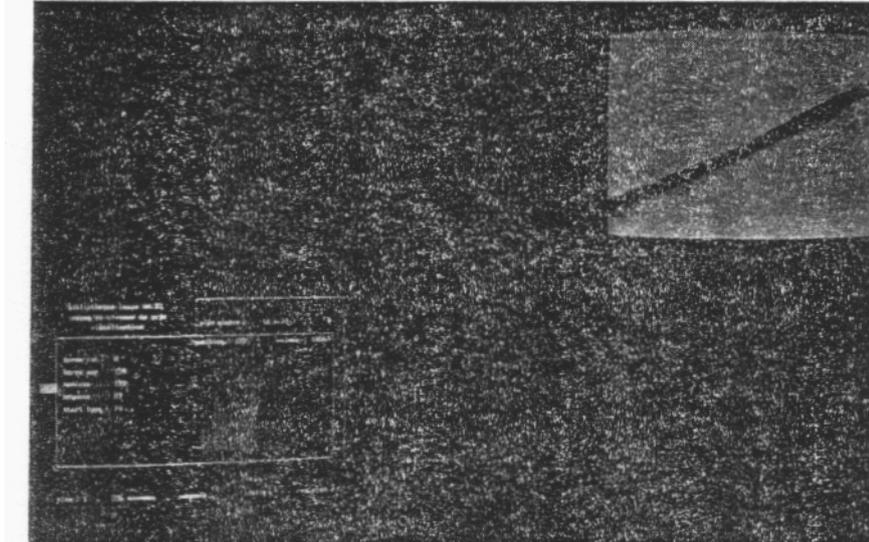


Figura 4

(a) Imagem de um fio têxtil sem pilosidade a (b) a correspondente distribuição de níveis de cinzento. Entropia: 0.678.

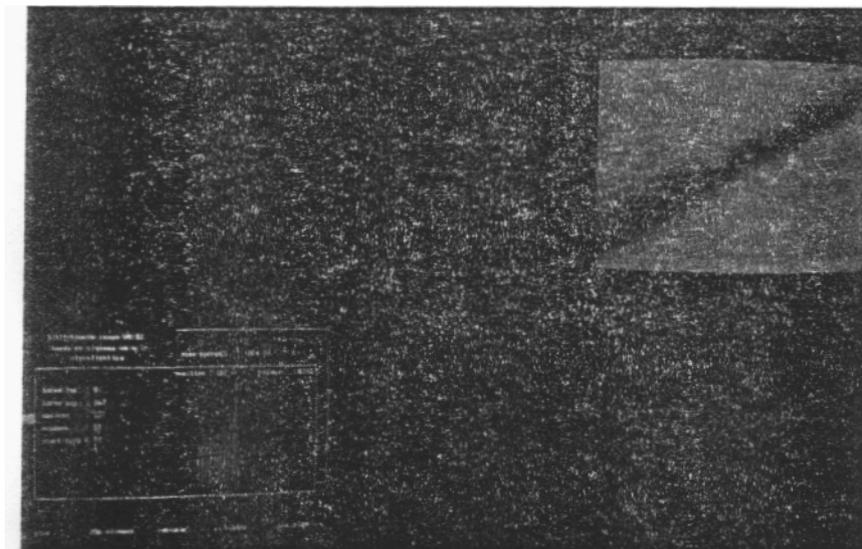


Figura 5

(a) O mesmo fio da figura 4 com pilosidade a (b) a respectiva distribuição de níveis de cinzento. Entropia: 0.7069.

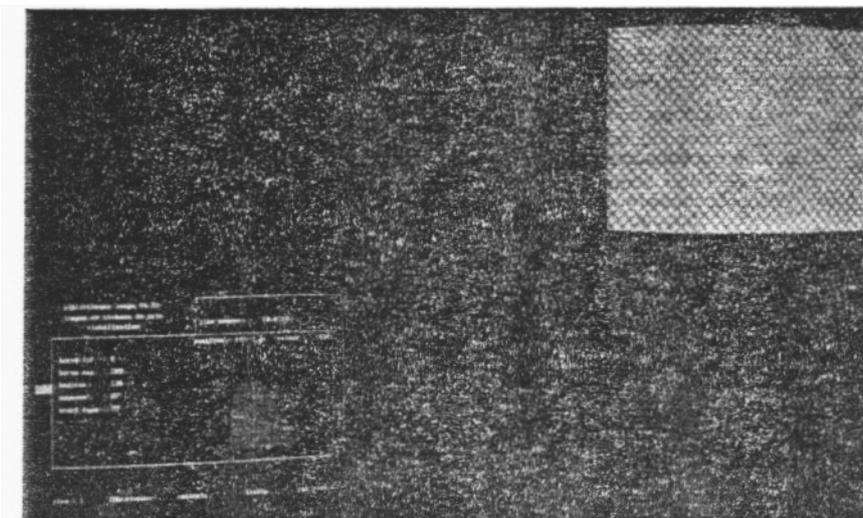


Figura 6

(a) Imagem de uma malha sem defeitos a (b) a respectiva distribuição de níveis de cinzento. Entropia: 3.1673.

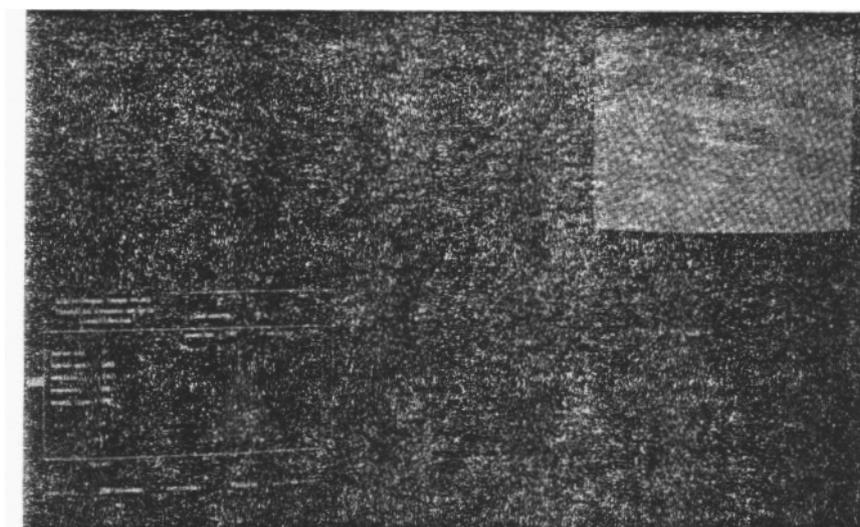


Figura 7

(a) A mesma malha que na figura 6 mas com defeitos a (b) a respectiva distribuição de níveis de cinzento. Entropia: 3.2806.

entrópica tem uma realidade palpável. Com a energia gasta na penteação almeja-se tornar um conjunto de fibras orientadas caoticamente num conjunto de fibras paralelizadas. Pretende-se com o processo penteação diminuir a entropia do sistema.

O resultado do seguimento do valor entrópico da lã penteada, utilizando o método descrito, ao longo das etapas do processo mostra claramente uma evolução no sentido de uma maior ordem no sistema com a consequente diminuição da entropia da informação.

4.1.2. Medidas de pilosidade de fios têxteis

A caracterização do grau de pilosidade de fios é um dos exemplos mais paradigmáticos da capacidade da utilização do método de medida de entropia aqui descrito.

Para os dois casos expressos nas figuras 4 e 5 a variação de entropia de informação é no sentido que a teoria prevê. A uma maior pilosidade associa-se uma maior desordem a que corresponde uma maior entropia.

4.1.3. Caracterização de defeitos em tecidos

A determinação de forma automática de defeitos em tecidos constitui uma tarefa de duplicação de uma actividade humana que melhor se pode descrever como sabedoria. Recorrer à ciência para resolver tais casos traduz-se, quase sempre, por uma grande limitação de resultados. A detecção de defeitos em tecidos é um destes casos.

Parece, no entanto, que a caracterização de tecidos por um parâmetro de medida da informação constituirá um avanço no sentido da tomada de decisão baseada em apenas um parâmetro, a entropia da informação.

Assim, no domínio da informação um defeito é qualquer variação da entropia da informação do tecido relativamente ao valor que é medido para o tecido sem defeitos.

Os exemplos ilustrados nas figuras 6 e 7 representam casos típicos a concludentes.

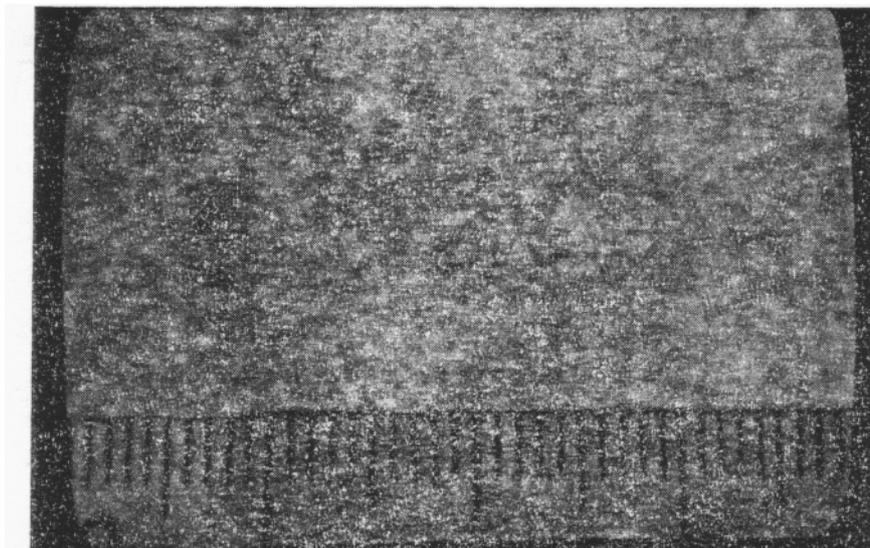


Figura 8

Imagem de papel Kraft. Entropia: 2.243.

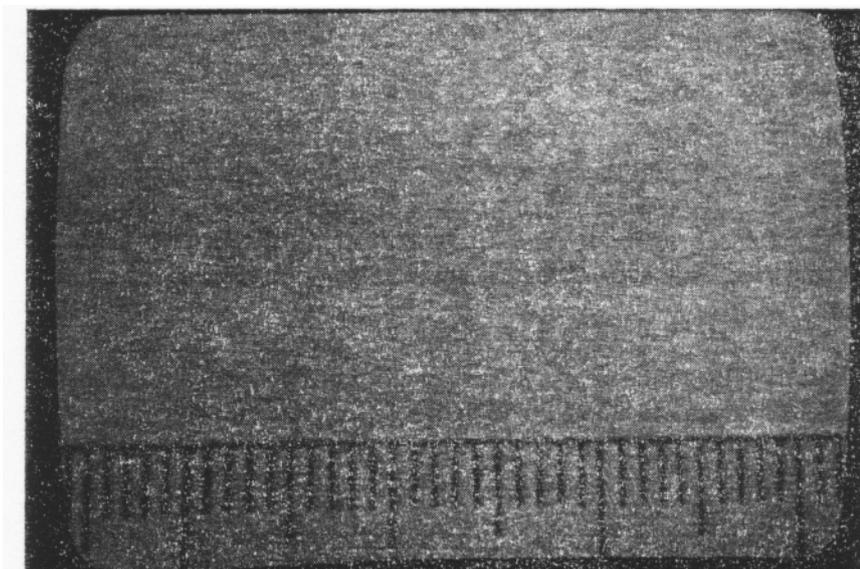


Figura 9

Imagem de papel para impressão. Entropia: 2.649.

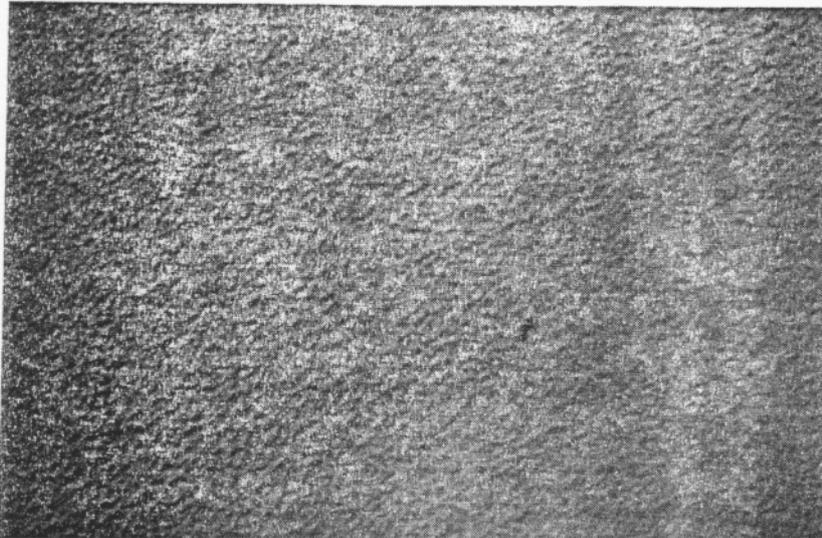


Figura 10
Cabedal sem defeitos. Entropia: 6.284.

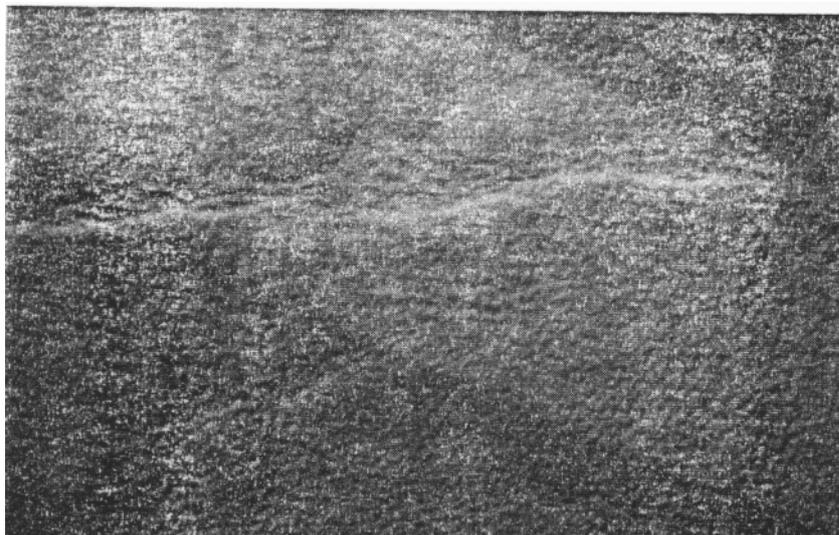


Figura 11
Cabedal com defeitos. Entropia: 5.462.

4.2. Sector do Papel

A distribuição espacial das fibras no papel dita um conjunto de propriedades físicas e químicas determinantes do aproveitamento que dele se faz.

Antes da possibilidade de determinação do grau de desordem na distribuição espacial das fibras que a medida da entropia da informação trouxe a medida da heterogeneidade na distribuição requeria métodos de difracção óptica. Estas implicam a utilização de lasers, que sempre têm uma vida limitada, e o tratamento em paralelo de grande número de dados uma vez que a medida têm que ser feitos numa área suficientemente extensa.

A diferença em entropia da informação mostra, para o exemplo da figura 6 claramente uma grande diferença caso se trate de um papel Kraft ou papel normal.

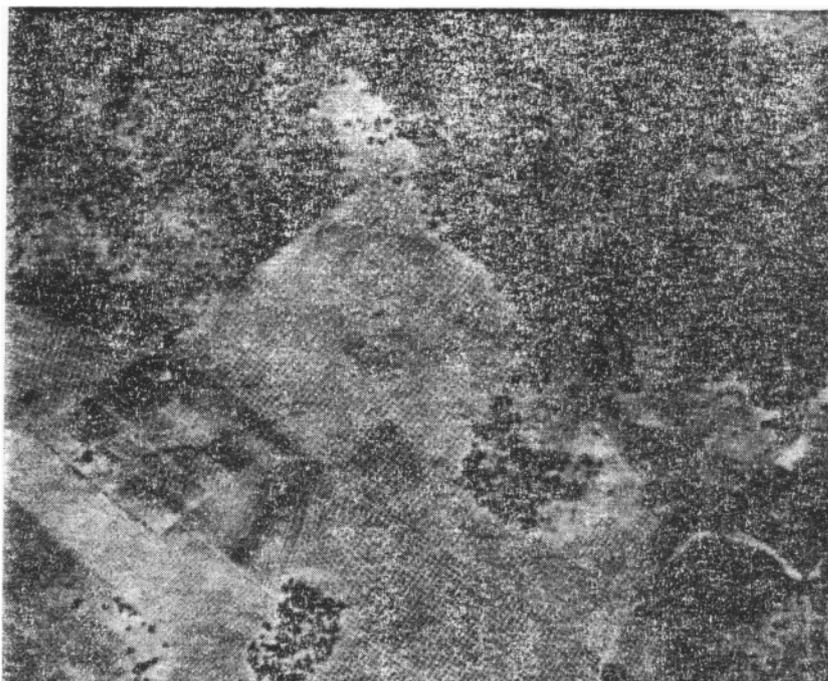


Figura 12

Fotografia aérea da região agrícola com sobreiros e oliveiras.

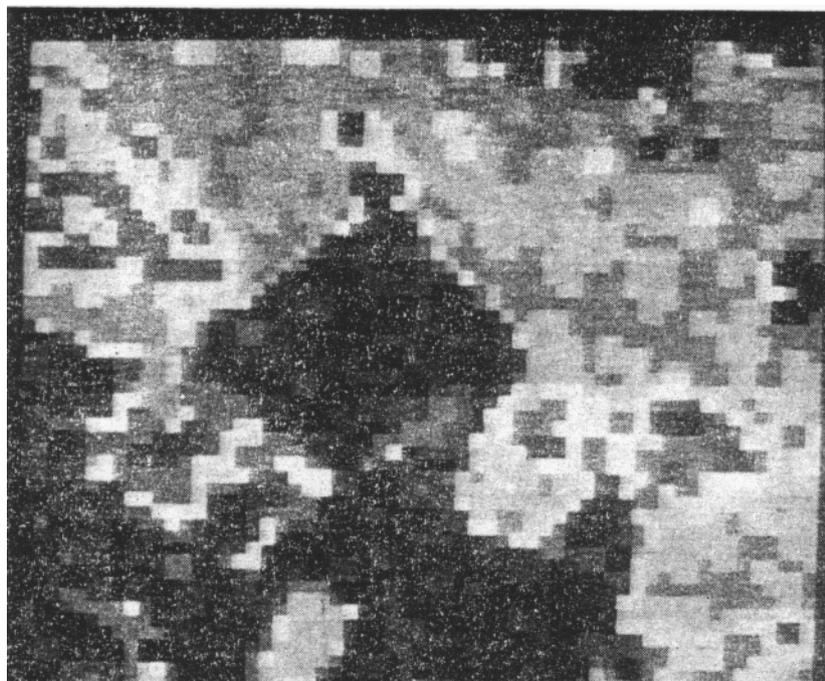


Figura 13
Mapa entrópico da distribuição das oliveiras a sobreiros.

O primeiro corresponde a uma distribuição espacial mais caótica enquanto que o segundo apresenta uma distribuição espacial bastante mais ordenada uma vez que este tipo de papel a mercê do processo de fabrico as fibras tendem a organizar-se segundo uma direcção preferencial.

4.3. Sector dos Curtumes

Mais uma vez a caracterização de uma pele a consequente monitoração de desvios ou defeitos não é um problema a que a ciência tenha dado respeito cabal.

A medida da entropia da informação permite, contudo, ter um método para a avaliação de desvios da pele relativamente como que for julgado como apropriado,. Qualquer desvio em relação ao

valor da entropia da informação considerado como sendo o caracterizador da pele sem defeitos é indicador da presença de desvios em relação à pele considerada, para um determinado fim, sem defeitos.

4.4. Cartografia

As espécies agrícolas quando localizadas numa forma densa no terreno apresentam características de distribuição espacial que são peculiares. É este o caso de oliveiras que ocorrem, muitas vezes, em espaços onde há outras espécies agrícolas nomeadamente sobreiros.

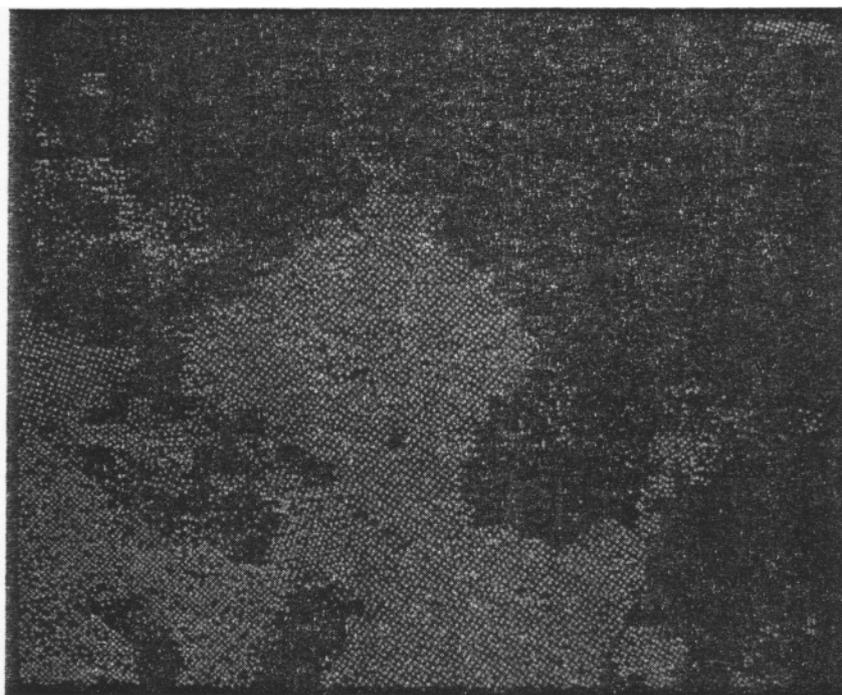


Figura 14
Imagem obtida da figura 12 depois de filtrada
pela máscara entrópica da figura 13.

Da figura 12 verifica-se que há mais ordem posta na plantação das oliveiras do que na disseminação das restantes espécies agrícolas. Por divisão em janelas sobre a imagem (figura 12) pode constituir-se um mapa de medidas de entropia da informação daquela região agrícola (figura 13).

Fazendo corresponder acima de uma entropia de informação de 4,5 bits a zona não ocupada por oliveiras contrói-se a partir da fig. 13 uma máscara que deixa ficar na imagem apenas as oliveiras.

A sua contagem a partir desta fase é imediata. Na figura 14 apresenta-se o resultado da operação da utilização da entropia de informação de uma fotografia aérea como discriminante para separar de forma expedita duas espécies agrícolas a proceder à respectiva cartografia a contagem. A contagem das oliveiras (1761) verificou-se ter apenas um desvio- de cinco por cento.

5. CONCLUSÕES

A entropia de informação é uma grandeza de fácil medida na zona do espectro electromagnético quando é possível organizar a informação numa imagem.

A utilização desta grandeza como discriminante revelou-se eficaz na resolução de um conjunto de problemas, sem solução anterior, em domínios tão complexos e díspares como são o têxtil, o papel, os curtumes e a cartografia temática de espécies agrícolas.

AGRADECIMENTOS

Ao Sr. Prof. Pinto Peixoto o acompanhamento participativo nos trabalhos. Ao Sr. Eng. Fernando Carvalho as discussões, a construção do equipamento optoelectrónico e a implementação dos conceitos. Ao Sr. Prof. Santos Silva todos os dados e os materiais relativos à aplicação dos conceitos e da tecnologia a problemas do sector têxtil, dos curtumes e do papel. Ao Sr. Doutor José Rebordão os resultados obtidos no âmbito do projecto SATCART no domínio da cartografia temática a partir de fotografia aérea.

REFERENCIAS

- SHANON, E. C., WEAVER, W., *The mathematical theory of communications*, University of Illinois Press, Chicago, 1949.
- PEIXOTO, J. P., *Introdução à teoria da informação*, Faculdade de Ciências de Lisboa, 1967.
- CARVALHO RODRIGUES et al., *The hidden parameter in textile processes: informational entropy*, J. Text. Institute, 1989, 80-n.º 4, Textile Institute.
- CARVALHO RODRIGUES et al., *Application of vision measurement in fabric inspection*, Tex. Res. J. (aceite para publicação em 1991).
- CARVALHO RODRIGUES et al., *Mesurement of randomness in fiber distribution in paper using computer vision*, TAPPI-Journal, December 1990.
- CARVALHO RODRIGUES et al., *Application of information entropy to defect characterisation in leather*, J. Soc. Leather Tech. and Chem., Dec. 1990.
(Comunicação apresentada à Classe de Ciências, na sessão de 1 de Fevereiro de 1990).