

FUSÃO DE IMAGEM PARA DETEÇÃO DE PESSOAS CAMUFLADAS

Nádia Bento ^{a1}, José Silva ^{a2}, José Dias ^b

^a CINAMIL - Centro de Investigação da Academia Militar, Academia Militar, Lisboa, Portugal

^b Departamento de Engenharia Electrotécnica e de Computadores, Instituto Superior Técnico, Lisboa, Portugal

RESUMO

A utilização de imagens térmicas constitui um benefício para as Forças Armadas. Devido às suas inúmeras vantagens, têm um grande número de aplicações, entre as quais a deteção de pessoas camufladas. Para melhorar a visualização, a imagem visível pode ser combinada com a imagem térmica. A fusão de imagem gerada por múltiplas fontes permite um maior detalhe, resultando num maior grau de segurança. No presente trabalho foram implementados três métodos de fusão de imagem ao nível do pixel, baseados em: análise em componentes principais, pirâmide laplaciana e transformada wavelet discreta. A análise qualitativa dos resultados permitiu concluir que melhor método é aquele que utiliza as wavelets, seguido do método que utiliza a pirâmide laplaciana. De seguida, foi efetuada uma análise quantitativa usando como métricas de desempenho o desvio padrão, a entropia, a frequência espacial, a informação mútua, o índice de qualidade de fusão e o índice de semelhança estrutural. Os valores obtidos sustentam as conclusões extraídas da análise qualitativa.

ABSTRACT

The use of thermal imaging is a benefit for the Armed Forces. Due to its numerous advantages, these images have a large number of applications, including the detection camouflaged people. To improve the image visualization, the visible image can be combined with the thermal image. This fusion produced by multiple sources allows a better detail, resulting in a higher degree of security. In this work three image fusion methods were implemented, based on: principal component analysis, Laplacian pyramid and discrete wavelet transform. The qualitative analysis of the results showed

¹ O presente artigo resultou do trabalho desenvolvido para a Dissertação de Mestrado elaborada pela autora.

² Email para contacto: jose.silva@academiamilitar.pt

that the best method is the one that uses wavelets, followed by the method using the Laplacian pyramid.

A quantitative analysis was also performed using several performance metrics: the standard deviation, entropy, spatial frequency, mutual information, the quality index of merged image and the structural similarity index. The results achieved support the conclusions extracted from the qualitative analysis.

Palavras chave: fusão de imagem, wavelets, pirâmide laplaciana, análise das componentes principais, métrica de desempenho, segurança.

1. INTRODUÇÃO

A maioria dos atuais sistemas de vigilância utiliza sistemas de detecção que operam apenas no espectro visível, o que constitui uma limitação. Por este motivo, tem sido proposto o uso de câmaras operando no espectro infravermelho-térmico para capturar a imagem térmica dos objetos. Desta forma, a utilização de imagens térmicas tem vindo a constituir um benefício para as Forças Armadas, devido à sua capacidade de utilização diurna e noturna e em diversas condições atmosféricas. Neste contexto, estas imagens podem ser usadas para a detecção de pessoas camufladas.

Para melhorar os resultados da detecção de pessoas, a informação capturada pelas câmaras do espectro visível pode ser combinada com a informação capturada pelas câmaras do espectro infravermelho-térmico. A fusão das duas imagens fornece uma boa perceção visual e permite uma melhor detecção de pessoas. Assim, o objetivo do trabalho realizado foi a implementação de métodos de fusão de imagem ao nível do pixel para detetar pessoas camufladas, ou seja, pessoas que se encontrem dissimuladas no terreno e que não são passíveis de serem observadas a olho nu, mantendo o enquadramento da paisagem envolvente.

2. ESTADO DA ARTE

No contexto da crescente procura por uma maior segurança e vigilância, tornam-se necessários sistemas de vigilância mais sofisticados e robustos. Naidu e Raol (2008) fazem uma comparação da fusão ao nível do pixel utilizando wavelets e análise em componentes principais (PCA). As métricas do desvio padrão, entropia cruzada e frequência espacial foram consideradas as adequadas quando não existe imagem de referência. A fusão de imagem utilizando wavelets com um maior grau de decomposição possui melhor desempenho.

Zheng (2007) faz uma comparação de algoritmos de fusão multi-escala ao nível do pixel, nomeadamente: baseados em pirâmides, discrete wavelet transform (DWT) e DWT Iterativa. Zheng propôs o algoritmo Advanced Discrete Wavelet Transform para a fusão de imagem. Os resultados quantitativos mostram que os algoritmos iterativos apresentam melhores desempenhos, seguidos da pirâmide laplaciana e por último a DWT.

Sadhasivam, et al., (2011) implementam um algoritmo que utiliza a DWT em conjunto com PCA. O coeficiente de baixa frequência é escolhido de acordo com a regra do máximo. Os pesos para a fusão são determinados por aplicação da PCA aos coeficientes de alta frequência. A semelhança estrutural e a informação mútua na imagem fundida apresentam melhores resultados para o método proposto.

Zheng, et al., (2011) desenvolvem um algoritmo que incorpora PCA na DWT. A PCA é aplicada aos coeficientes de aproximação, ou seja, os coeficientes de baixa frequência, enquanto os coeficientes de detalhe são escolhidos de acordo com o maior valor absoluto. O algoritmo proposto é comparado com outras técnicas de fusão utilizando a entropia, a frequência espacial e o índice de qualidade de imagem, para os casos em que não existe imagem de referência.

3. MÉTODOS

Nesta secção descrevem-se os métodos utilizados: PCA, pirâmide laplaciana e wavelets, e ainda são apresentadas as métricas para a análise quantitativa.

3.1. ANÁLISE EM COMPONENTES PRINCIPAIS

A fusão de imagem baseada em PCA é alcançada através de uma média ponderada das imagens a serem fundidas. Os pesos para cada imagem fonte (imagem obtida por uma câmara) são obtidos a partir dos vetores próprios correspondentes ao maior valor próprio das matrizes de covariância de cada imagem fonte, isto é, as componentes principais. A imagem fundida é determinada por (Naidu, 2008):

$$I_{fund} = p_1 I_{vis} + p_2 I_{term}, \quad (1)$$

I_{vis} e I_{term} denotam as imagens visível e térmica, $p_1 = \frac{v_{1i}}{v_{1i}+v_{2i}}$ e $p_2 = \frac{v_{2i}}{v_{1i}+v_{2i}}$ são calculados a partir da matriz de covariância $\mathbf{c} = \mathbf{v} \Sigma \mathbf{v}^t$, em que \mathbf{c} , \mathbf{v} , Σ são

matrizes de dimensão 2×2 , $\mathbf{v} = [v_1, v_2]$ e $\Sigma = \text{diag}(\lambda_1, \lambda_2)$. O par $(\mathbf{v}_i, \lambda_i)$ com $i = 1, 2$ corresponde ao par vetor próprio/valor próprio de c , em que $\mathbf{v}_i = \begin{bmatrix} v_{1i} \\ v_{2i} \end{bmatrix}$. Foram implementadas três variantes deste método, sendo estes designadas por PCA1, PCA2 e PCA3.

3.2. PIRÂMIDE LAPLACIANA

Neste método, a imagem fonte é decomposta em subimagens com diferentes resoluções espaciais. A fusão de imagem baseada em pirâmides consiste em usar uma regra de fusão para construir uma representação de uma pirâmide fundida a partir das representações em pirâmide das imagens fonte. A imagem fundida é obtida ao fazer a transformada inversa da pirâmide. Na pirâmide laplaciana, o nível mais baixo da pirâmide é construído a partir da imagem original e cada um dos restantes níveis é construído recursivamente a partir do seu nível mais baixo, aplicando quatro operações básicas [5]: *blurring*, sub amostragem, expansão por interpolação e diferenciação.

De forma a derivar a pirâmide laplaciana, existem duas operações padrão: “Reduzir” e “Expandir” (inverso da “Reduzir”), cujo objetivo é expandir um array $(M+1) \times (N+1)$ num array $(2M+1) \times (2N+1)$, ao preencher M zeros na direção horizontal e N zeros na direção vertical. Desta forma, “Expandir” aplicado ao array X_l da pirâmide gaussiana gera um array \hat{X}_l com o mesmo tamanho que X_{l-1} .

Consideram-se duas operações padrão:

$$X_l = \text{Reduzir}(X_{l-1}) \quad (2)$$

$$\hat{X}_l = \text{Expandir}(X_l) \quad (3)$$

A pirâmide laplaciana é aplicada para decompor as imagens fonte em N níveis:

$$\begin{cases} LP_l = X_l - \hat{X}_{l+1}, 0 \leq l < N \\ LP_N = X_N, l = N. \end{cases} \quad (4)$$

A reconstrução da imagem a partir da pirâmide laplaciana é o processo inverso da decomposição e na direção inversa, ou seja, do nível do topo para o nível da base, com a seguinte definição:

$$\begin{cases} X_N = LP_N, l = N \\ X_l = LP_l + \hat{X}_{l+1}, 0 \leq l < N \end{cases} \quad (5)$$

Existindo duas imagens para fundir I_{vis} e I_{term} , a construção da pirâmide é feita para cada imagem individualmente.

3.3. WAVELETS

A representação em wavelets é muito utilizada em processamento de imagem e fornece uma decomposição multi-resolução de uma imagem. As imagens fonte são decompostas em coeficientes de aproximação e de detalhe, num determinado nível usando a DWT. Depois, os coeficientes de aproximação e de detalhe são combinados usando a regra de fusão \emptyset ; ao utilizar a DWT Inversa é gerada a imagem fundida:

$$I_{fund} = IDWT [\phi\{DWT(I_{vis}), DWT(I_{term})\}], \quad (6)$$

onde DWT e IDWT denotam, respectivamente, as transformadas DWT e IDWT.

As regras de fusão utilizadas neste trabalho foram: (i) o máximo dos coeficientes de aproximação à maior escala e o maior valor absoluto dos coeficientes de detalhe em cada escala da transformada; (ii) a média dos coeficientes de aproximação à maior escala e o maior valor absoluto dos coeficientes de detalhe em cada escala da transformada.

3.4. AVALIAÇÃO QUANTITATIVA

Esta secção apresenta uma avaliação qualitativa e quantitativa comparando o desempenho dos algoritmos de fusão descritos anteriormente. As métricas utilizadas são as seguintes:

O **Desvio Padrão** mede o contraste da imagem fundida; uma imagem com um elevado contraste terá um elevado valor de desvio padrão (Naidu, 2008). A **Entropia** mede o conteúdo de informação de uma imagem. Uma imagem com elevado conteúdo de informação terá uma entropia elevada (Eskicioglu, 1995). A **Frequência Espacial** indica o nível de atividade global de uma imagem, quanto maior for o seu valor mais informação possui a imagem fundida (Wang, 2004). A **Informação Mútua** mede o grau de dependência entre duas imagens. É calculada ao definir o histograma conjunto das imagens fonte e a imagem fundida. Uma dependência maior, ou seja, um valor maior, implica melhor qualidade (Sadhasivam, 2011). O **Índice de Qualidade de Fusão** de uma imagem mede a semelhança entre a imagem fundida e as imagens fonte. Esta métrica pode assumir valores entre 0 e 1, onde o valor 1 corresponde a uma melhor qualidade da imagem fundida (Zheng, 2004).

O **Índice de Semelhança Estrutural** indica a semelhança da informação estrutural entre duas imagens. Um valor elevado indica que a informação presente em cada uma das imagens fonte se encontra presente na imagem fundida (Zheng, 2004).

4. RESULTADOS

Nesta secção são apresentados os resultados para cada um dos métodos utilizados, seguidos de uma análise qualitativa e quantitativa.

4.1. CONJUNTO DE DADOS

As imagens utilizadas para o presente trabalho possuem 320×240 pixéis, obtidas com uma câmara FLIR T440bx em dois exercícios de campo da Academia Militar. Foram seleccionadas 20 imagens duplas (região do visível e região térmica) mais ilustrativas da diversidade de camuflagem natural ou artificial. Os algoritmos foram implementados em MATLAB®.

4.2. ANÁLISE QUALITATIVA

De forma a comparar os resultados dos três métodos implementados, a terceira variante da PCA, a pirâmide laplaciana com quatro níveis de decomposição e as wavelets com um nível de decomposição usando a 1ª regra de fusão foram aplicados aos 20 pares de imagens obtendo-se assim 20 imagens fundidas. Neste primeiro exemplo, ilustrado na Figura 1, a qualidade das duas primeiras imagens é semelhante. Na terceira imagem, a pessoa encontra-se mais destacada em relação às outras duas, correspondendo, por esse motivo, à melhor fusão. O detalhe presente nas três imagens é idêntico.

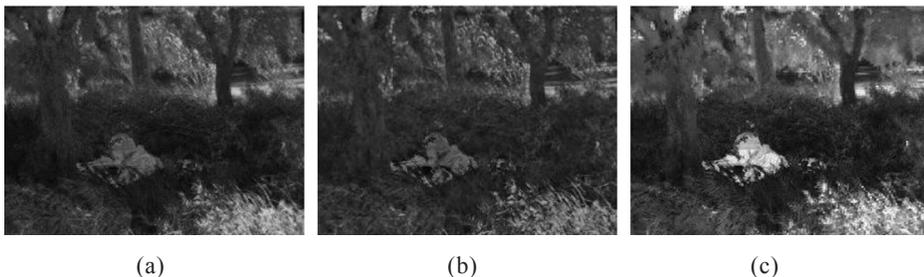


Figura 1 - a) imagem fundida com PCA b) imagem fundida com pirâmide laplaciana c) imagem fundida com wavelets

No segundo exemplo, ilustrado na Figura 2, encontra-se uma pessoa atrás de arbustos, o que por si só contribui para a falta de detalhe, pois os arbustos são esparsos e nas imagens fonte algumas zonas do corpo humano aparentam ser apenas uma mancha. A segunda imagem, obtida com a pirâmide laplaciana, é a que aparenta o melhor compromisso entre o detalhe dos arbustos e a componente térmica da pessoa camuflada. Na segunda imagem, é identificada não só a pessoa, mas também perceptível o ambiente envolvente.

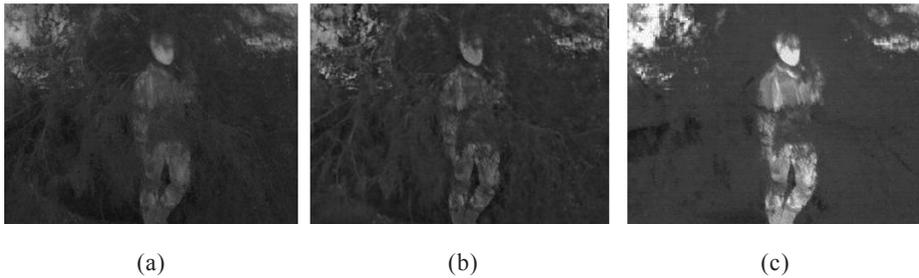


Figura 2 - a) imagem fundida com PCA b) imagem fundida com pirâmide laplaciana c) imagem fundida com wavelets

No terceiro exemplo, ilustrado na Figura 3, a qualidade subjetiva das três imagens é idêntica. Porém, a terceira imagem mostra a componente térmica de uma forma mais clara, correspondendo, por isso, à melhor fusão.

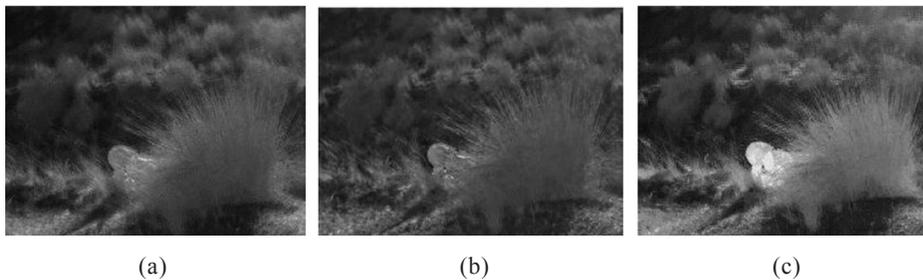


Figura 3 - a) imagem fundida com PCA b) imagem fundida com pirâmide laplaciana c) imagem fundida com wavelets

A partir da análise qualitativa, o método baseado em wavelets aparenta produzir os melhores, seguido do método que utiliza a pirâmide laplaciana. Um dos fatores que contribuí para os fracos resultados apresentados pelo método que utiliza a PCA é o facto de quando ambos os valores das

componentes principais são próximos de 0.5, a fusão assemelha-se à fusão através da média simples o que produz um contraste baixo. Por outro lado, o método que utiliza as wavelets é o que apresenta melhores resultados, pois é aquele que faz a fusão das várias componentes da imagem.

4.2. ANÁLISE QUANTITATIVA

Nesta secção é efetuada uma avaliação quantitativa dos métodos utilizados na fusão de imagem para uma comparação objetiva e fundamentada do desempenho de cada método. Os resultados serão apresentados em gráficos (*diagrama de extremos e quartis*).

Nestas figuras seguintes, é feita uma comparação objetiva entre os vários métodos, em que cada coluna da *boxplot* corresponde a um método, sendo estes designados por PCA1, PCA2 e PCA3 correspondendo às três variantes da PCA implementadas; LP02 e LP04, correspondendo aos métodos que utilizam a pirâmide laplaciana com dois e quatro níveis; WV11 e WV12, correspondendo à wavelet com um nível de decomposição com a 1ª e 2ª regra de fusão; por último, WV21 e WV22, correspondendo à wavelet com dois níveis de decomposição com a 1ª e 2ª regra de fusão.

4.2.1 Informação Mútua

A informação mútua mede o grau de dependência entre duas imagens. O valor da informação mútua é a soma entre a informação mútua de cada imagem fonte com a imagem fundida, portanto quanto maior for o seu valor maior será a dependência entre as imagens fonte e a imagem fundida.

Os melhores resultados são obtidos quando se utiliza as wavelets com a 1ª regra fusão, tanto com um como com dois níveis de decomposição; os resultados com a PCA são os piores (ver figura 4). Os valores mais elevados devem-se ao facto de, ao serem seleccionados os pixéis com maior intensidade, a transferência de informação da imagem original é maior, resultando num melhor resultado. Por outro lado, na pirâmide laplaciana e nas wavelets com a 2ª regra de fusão é utilizada a média dos pixéis, e por esse motivo os valores da informação mútua são mais baixos.

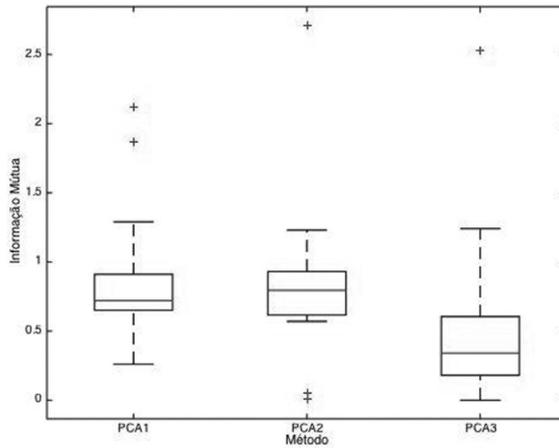


Figura 4 - Boxplot Informação Mútua

4.2.2 Índice de Qualidade de Fusão

O índice de Qualidade de Fusão mede a semelhança entre a imagem fundida e as imagens fonte. Possui valores entre 0 e 1, sendo que quanto maior for a semelhança mais próximo estará do valor 1. Idealmente, a imagem fundida deve apresentar ambas as imagens fonte.

Tal como se pode ver na figura 5, os melhores resultados são obtidos quando se utiliza o método com PCA, seguidos do método que utiliza a pirâmide laplaciana.

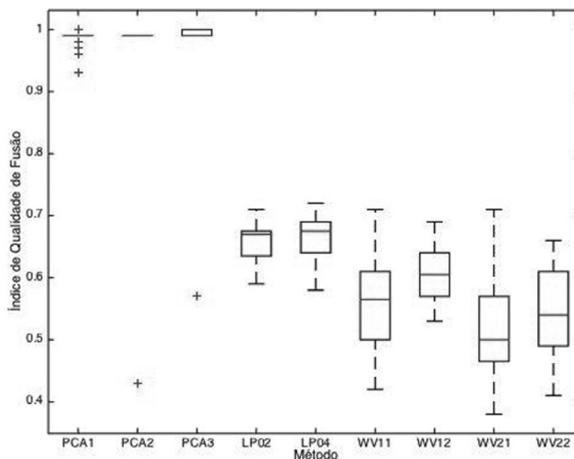


Figura 5 - Boxplot Índice de Qualidade de Fusão

Contrariamente ao que se tem verificado nas métricas anteriores, os métodos baseados na PCA obtém os valores mais elevados. Isto pode ser explicado pelo facto de que, neste método, é feita uma média ponderada das imagens fonte para obter a imagem fundida e assim, ambas as imagens fonte contribuem para a imagem final da mesma forma. Ou seja, enquanto nos métodos que utilizam as wavelets existem coeficientes que são seleccionados apenas de uma das imagens fonte (havendo coeficientes de detalhe de uma das imagens fonte que não contribuem para a imagem fundida), nos métodos baseados na PCA isso não acontece.

5. CONCLUSÕES

O presente trabalho surgiu da necessidade da implementação de um método que permita a deteção de pessoas camufladas, enquanto estas estão enquadradas no seu ambiente, tornando a sua camuflagem ineficiente. Os três métodos de fusão de imagem implementados foram: o método baseado em análise em componentes principais (PCA), o método que utiliza a pirâmide laplaciana e o método que usa a análise em wavelets.

Foi efetuada uma análise qualitativa das imagens fundidas baseada em conjuntos de três imagens. Desta análise, conclui-se que o método que produz melhores resultados é aquele que utiliza as wavelets, seguido do método que utiliza a pirâmide laplaciana e por último o que utiliza a PCA.

Na análise quantitativa baseada em seis métricas de desempenho, concluiu-se que o desvio padrão, a entropia e a frequência espacial são métricas apropriadas para fazer uma comparação relativa entre os métodos implementados (pois medem a qualidade da imagem fundida), contudo, por si só estas métricas não têm em consideração a relação existente entre a imagem fundida e as imagens fonte. Por outro lado, a informação mútua, o índice de qualidade de fusão e o índice de semelhança estrutural têm presente aquela relação, sendo por isso consideradas as métricas mais adequadas para medir a qualidade da fusão de imagem.

Ao comparar os resultados obtidos qualitativa e quantitativamente, observa-se que os métodos tidos como melhores na avaliação qualitativa são de facto os melhores na avaliação quantitativa, ou seja, o método que utiliza as wavelets com um nível de decomposição com a primeira regra de fusão e o método que usa a pirâmide laplaciana com quatro níveis.

REFERENCIAS

Naidu, V. P. S., Raol, J. R. (2008): Pixel-Level Image Fusion using Wavelets and Principal Component Analysis. *Defence Science Journal* 58, pp. 338-352

Zheng, Y.: Multi-scale Fusion Algorithm Comparisons: Pyramid, DWT and Iterative DWT. In *12th International Conference on Information Fusion*, pp. 1060-1067 (2009)

Sadhasivam, S. K., Keerthivasan, M. B., Muttan, S.: Implementation of Max Principle with PCA in Image Fusion for Surveillance and Navigation Application. In *Electronic Letters on Computer Vision and Image Analysis*, Universitat Autònoma de Barcelona, pp. 1577-5097 (2011)

Zheng, Y., Essock, E. A., Hansen, B. C.: An Advanced Image Fusion Algorithm Based on Wavelet Transform -Incorporation with PCA and Morphological Processing. University of Louisville (2004)

Leung, L. W., King, B., Vohora, V.: Comparison Of Image Data Fusion Techniques Using Entropy and INI. In *Centre for Remote Imaging, Sensing and Processing*, National University of Singapore (2001)

Eskicioglu, A. M., Fisher, P. S.: Image Quality Measures and Their Performance. *IEEE Transactions On Communications* 43, 2959-2965 (1995)

Wang, Z., Bovik, A. C., Sheikh, H. R., et al.: Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Transactions On Image Processing* 13, 1-14, (2004)

Zheng, Y., Essock, E. A., Hansen, B. C., et. al.: A new metric based on extended spatial frequency and its application to DWT based fusion algorithms. *Information Fusion* 8, pp. 177-192 (2007)

Nádia Bento é Tenente de Transmissões e Mestre em Engenharia Eletrotécnica Militar.

José Silva é Professor Auxiliar do Departamento de Ciências e Tecnologias de Engenharia, Academia Militar. Desenvolve as suas atividades de investigação no Centro de Investigação da Academia Militar.

José Dias é Professor Associado do Departamento de Engenharia Eletrotécnica e de Computadores, Instituto Superior Técnico. Desenvolve as suas atividades de investigação no Instituto de Telecomunicações, ramo de Lisboa - Instituto Superior Técnico.

