

Goodears – Robot Orientado por Farol Acústico

José Vieira, Luís Almeida, Paulo Santos, João Reis
vieira@det.ua.pt, lda@det.ua.pt, athos@portugalmail.pt, joao.m.rei@portugalmail.pt

DET - IEETA
Universidade de Aveiro
P-3810-193 Aveiro, Portugal

Resumo

Este artigo descreve um pequeno robot capaz de seguir uma fonte sonora apenas com dois microfones. No caso de uma fonte conhecida utilizou-se um sistema de detecção baseado em “matched filter” que se revelou muito eficaz e imune ao ruído. No caso de uma fonte sonora desconhecida o método utilizado foi o da correlação cruzada entre os dois canais. Em ambos casos foi utilizado o atraso entre os canais para determinar a direcção da fonte sonora. Verificou-se igualmente que o robot era capaz de descobrir a fonte mesmo quando esta se apresenta na retaguarda.

Palavras chave: detecção de fonte sonora, robot autónomo.

1. Introdução

A determinação da localização de uma fonte sonora é de grande importância para a maior parte dos mamíferos, tendo o sistema auditivo evoluído de forma a proporcionar rapidez e precisão na localização. Esta capacidade é utilizada para identificar pontos de interesse no meio circundante e direccionar o sistema visual, que possui uma maior resolução apenas na parte central do campo de visão. Esta característica é igualmente importante nos seres humanos e pode ser determinante para um robot autónomo que tenha de partilhar o espaço com seres humanos [1]. A audição pode ser um sentido importante num robot autónomo, dado que pode complementar outras formas de percepção do meio como é o caso da visão. Existem igualmente outras aplicações para um sistema capaz de localizar a posição de uma fonte sonora: direccionamento das câmaras de vídeo em sistemas de vídeo-conferência, sistemas de vigilância, etc.

Grande parte dos sistemas de detecção da posição de uma fonte sonora recorrem a matrizes de sensores e a técnicas de processamento de sinal intensivo que limitam muito a sua aplicação prática. Alguns investigadores [1,5] têm desenvolvido técnicas inspiradas no sistema auditivo dos mamíferos uma vez que estes conseguem bons resultados apenas com dois sensores. Para determinação do azimute o sistema auditivo humano utiliza duas técnicas [2]: o atraso de um sinal relativamente a cada ouvido (ITD – Interaural Time Difference) e a diferença de intensidade de um sinal em cada um dos ouvidos (IID – Interaural Intensity Difference). Para determinar a elevação de uma fonte sonora o sistema auditivo humano tira partido da forma do pavilhão auditivo, que provoca atenuação em frequências que dependem da elevação da fonte [2]. Inspirado nesta “codificação espectral” da direcção de uma fonte acústica, é possível com apenas

um microfone munido de um reflector construir um sistema que determina a direcção de uma fonte sonora [5].

No caso do robot construído, é apenas utilizado o atraso na chegada de uma frente de onda a cada um dos microfones. Consideraram-se duas situações para o problema da localização da fonte sonora: primeiro uma fonte sonora conhecida e em segundo uma fonte sonora desconhecida. O primeiro problema é bastante mais simples uma vez que se escolheu para fonte sonora (farol) um sinal que fosse facilmente detectado, neste caso “chirps” transmitidos periodicamente. O segundo problema é bastante mais complexo, especialmente em ambientes reverberantes [2], tendo sido implementado apenas um simples algoritmo de correlação cruzada entre o canal esquerdo e o direito. Na figura 1 podemos observar a geometria do problema de localização da fonte. Os microfones estão separados de uma distância D_M e como se pode observar na figura para distâncias D_L e D_R (de “left” e “right”) muito maiores que o espaçamento D_M entre microfones, o ângulo θ a que está a fonte pode ser aproximado por

$$q \cong \sin^{-1}\left(\frac{D_L - D_R}{D_M}\right) \quad (1)$$

ou em termos da ITD

$$q \cong \sin^{-1}\left(\frac{v_s ITD}{D_M}\right) \quad (2)$$

em que v_s é a velocidade do som.

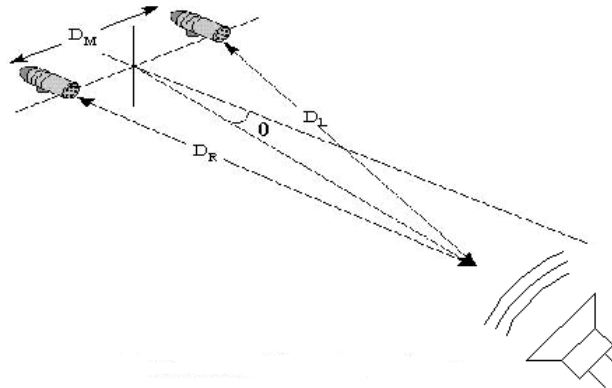


Fig. 1 – Visão geral do sistema com os dois microfones separados de uma distância D_M e a geometria de detecção da posição da fonte sonora.

O robot construído possui como unidade central de processamento um DSP C30 da Texas Instruments que para além de realizar todo o processamento de sinal controla ainda os motores eléctricos.

Na secção 2 faremos uma descrição detalhada do algoritmo de detecção do farol acústico baseada em “matched filters” e PLL’s por “software”. Na secção 3 faremos uma descrição do algoritmo de detecção de uma fonte sonora desconhecida pelo método da correlação cruzada. Na secção 4 será demonstrado que apesar do robot possuir apenas dois microfones, com um simples algoritmo de pesquisa da fonte sonora este acaba sempre por a detectar correctamente. Finalmente serão apresentadas as conclusões bem como alguns resultados preliminares.

2. Detecção de uma fonte sonora conhecida

Nesta situação o robot deve ser capaz de detectar a localização de um farol acústico que emite um sinal sonoro conhecido. Para tirar o máximo proveito deste facto escolheu-se para sinal emitido pelo farol um “chirp” na gama audível (300-3400Hz). A detecção deste sinal é realizada através de “matched filter” que permite obter boa imunidade ao ruído ambiente. Além disso, o “chirp” é enviado de forma periódica, permitindo através de uma PLL implementada por “software”, evitar detecções falsas. Na figura (2) podemos observar um diagrama de blocos do processamento realizado pelo DSP para a detecção da posição do farol acústico.

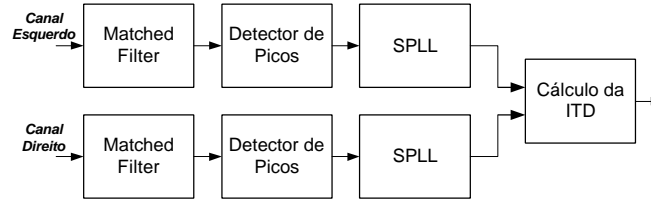


Fig. 2 - Diagrama de blocos do método de detecção do farol acústico.

2.1 Sinal enviado “chirp”

Este tipo de sinal pode ser descrito como uma sinusóide cuja frequência varia linearmente ao longo de um período T de tempo, desde uma frequência inicial f_0 até à frequência f_1 . Este sinal pode ser igualmente descrito como a modulação em frequência de um sinal em dente de serra, representado pela expressão

$$x(t) = A \cos(\mathbf{q}(t)), \quad (1)$$

em que $\mathbf{q}(t)$ é a fase instantânea dada pela seguinte expressão

$$\mathbf{q}(t) = 2pm \times t^2 + 2pf_0 t, \quad (2)$$

em que o declive do dente de serra é dado pela constante m , com valor dado pela expressão

$$m = \frac{f_1 - f_0}{2T}.$$

2.2. “Matched filter” e detector de picos

Consideremos um filtro digital do tipo FIR com resposta impulsional $h(n)$ com N amostras, sendo a resposta a um sinal $x(n)$ dada pela convolução de $x(n)$ com $h(n)$ tal como é indicado na seguinte expressão

$$y(n) = \sum_{k=0}^{N-1} h(k)x(n-k). \quad (3)$$

Vamos supor que pretendemos detectar um sinal $s(n)$ com N amostras, utilizando o filtro anterior. Que valor devem ter os coeficientes $h(n)$ do filtro de modo a que um simples detector de picos colocado na sua saída maximize o resultado. Há muito que se sabe que a resposta para este problema é um filtro com os coeficientes dados por

$$h(n) = s(N-1-n), \quad (4)$$

ou seja o sinal que se pretende detectar transposto no tempo [3]. Este filtro é óptimo no sentido em que maximiza a relação $\left(\frac{\mathbf{e}}{\mathbf{s}}\right)^2$ quando $n = N-1$, em que \mathbf{e} é a energia do

señal $s(n)$ e σ^2 é a potência de ruído aleatório captado. Substituindo a equação (4) em (3) obtemos a seguinte expressão

$$y(n) = \sum_{k=0}^{N-1} s(N-1-k)x(n-k), \quad (5)$$

que como se pode constatar realiza a autocorrelação do sinal $s(n)$.

2.3. Software PLL

Apesar da detecção dos “chirps” ser realizada com “matched filters” devido à reverberação acústica e a outras fontes de ruído (os motores do robot por exemplo) optou-se por transmitir os “chirps” de forma periódica. No receptor, construiu-se uma PLL [4] por software com uma gama de captura bastante estreita, que entra em “lock” quando detecta um número determinado de “chirps” consecutivos. Desta forma consegue-se uma imunidade a ruído bastante boa e o robot raramente detecta falsos positivos. Na figura 3 podemos observar um diagrama de blocos com a descrição do algoritmo utilizado na PLL.

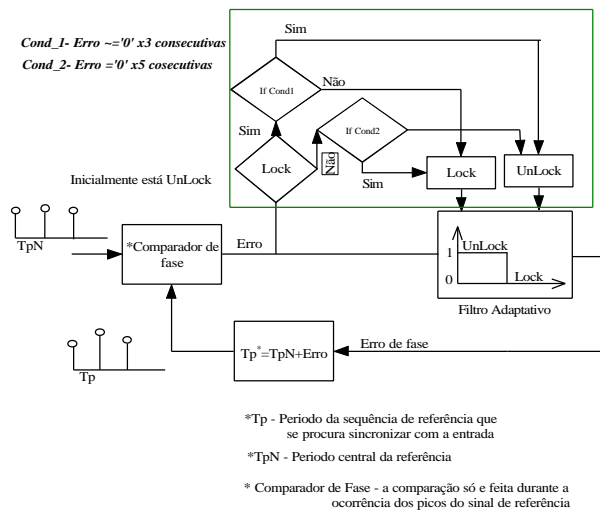


Figura 3. Diagrama de blocos da “software” PLL.

3. Detecção de uma fonte sonora desconhecida

Outro método para obter a direcção de uma fonte sonora é através da correlação cruzada dos sinais dos dois microfones. Para tal, usa-se uma janela de som stereo dos microfones e procura-se o máximo da função de correlação. A diferença entre o índice da amostra correspondente a $t=0$ menos o índice da amostra do máximo dá-nos o valor da ITD. O sinal, positivo ou negativo define a localização esquerda/direita no plano azimutal. O módulo da diferença, corresponde então ao atraso temporal entre os dois microfones. Contudo, o cálculo da correlação cruzada conforme expresso por

$$R_{xy}[t] = \sum_{n=0}^{N-1} x[n]y[n+t] \quad (6)$$

é muito pesado computacionalmente, em especialmente para janelas de dimensão elevada. Assim, para poder implementar o algoritmo da correlação em tempo real, torna-se necessário executá-lo amostra a amostra. Para tal aplicámos um algoritmo recursivo que se baseia num conceito elementar. Quando efectuamos a correlação cruzada seguinte, apenas deslocamos a janela (com N amostras) de uma amostra,

portanto, existem muitas operações em comum com a correlação anterior, aliás para cada n podemos subtrair o primeiro produto cruzado da correlação anterior e somar um produto cruzado novo tendo em conta as amostras recentes tal como se pode ver na equação seguinte

$$R_{xy}(n+1, t) = \begin{cases} R_{xy}(n, t) - x_{n-t}y_n + x_{n+N}y_{n+N+t} & t < 0 \\ R_{xy}(n, t) - x_ny_{n+t} + x_{n+N-t}y_{n+N} & t \geq 0 \end{cases} \quad (7)$$

Repare-se que neste algoritmo o cálculo da correlação no instante seguinte para um dado n não implica calcular a correlação para todos os t 's; podemos simplesmente limitar a gama aos atrasos possíveis. Por exemplo, se a ITD dá no máximo um desfaseamento de 8 amostras, não adianta calcular a correlação para $|t| > 8$, poupando assim muito tempo de processamento.

4. Algoritmo de orientação

A utilização de apenas dois microfones para determinar a direcção de uma fonte sonora com base na IDT não permite distinguir entre uma fonte localizada à frente e outra localizada atrás da linha que une os microfones. Na figura 4, podemos constatar que se um farol acústico F se encontrar localizado atrás da linha dos microfones o robot verá a imagem virtual I localizada à frente dessa linha. Contudo, como a informação azimutal obtida a partir da IDT é utilizada para orientação do robot, causando um determinado movimento angular e em direcção à fonte sonora detectada, o robot acaba sempre por se orientar para o farol acústico real. Na exposição que se segue considera-se que os microfones se encontram sobre a linha que une as rodas do robot, linha esta perpendicular à respectiva orientação R_k do robot no instante t_k .

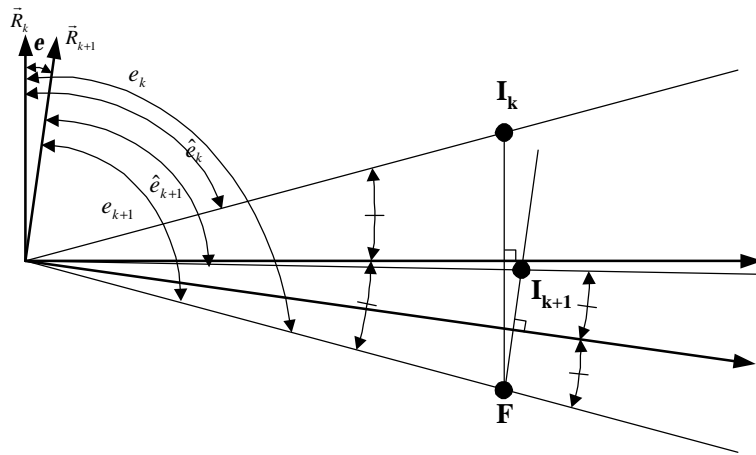


Fig. 4 – Geometria do problema de orientação do robot para o farol F .

Na figura 4 o robot, num dado instante t_k , está orientado na direcção R_k fazendo um ângulo e_k com o farol F . Como o farol se encontra atrás do robot, este tomará como posição do farol a imagem I_k e será para esta imagem que o robot se orientará rodando de um ângulo e . Pode-se concluir de imediato que após a rotação, o erro angular e_{k+1} em relação à verdadeira posição do farol F será diminuído de e ,

$$e_{k+1} = e_k - e. \quad (8)$$

Uma vez que o robot está a seguir a imagem \mathbf{I}_k do farol \mathbf{F} , coloca-se a questão de saber qual a nova posição desta \mathbf{I}_{k+1} depois da correcção angular \mathbf{e} . Da figura 4 podem-se deduzir as seguintes equações

$$e_k + \hat{e}_k = 180^\circ \quad (9)$$

$$e_{k+1} + \hat{e}_{k+1} = 180^\circ, \quad (10)$$

e das equações (8) e (10) podemos obter

$$\hat{e}_{k+1} = 180^\circ - e_k + \mathbf{e} \quad (11)$$

se substituirmos e_k na equação anterior a partir da equação (10) obtemos

$$\hat{e}_{k+1} = \hat{e}_k + \mathbf{e}, \quad (12)$$

que permite concluir que o erro em relação à imagem do farol aumentou de \mathbf{e} . Deste modo, o robot continuará a perseguir a imagem do farol até que o farol passe para a frente. Verifica-se assim que com um simples e intuitivo algoritmo de redução do erro em relação ao farol o robot acaba sempre por encontrar o farol mesmo que este se encontre inicialmente na retaguarda. Este comportamento foi verificado na prática.

5. Conclusões

No primeiro caso, a determinação da direcção foi efectuada apenas com recurso à diferença de fase entre os sinais recolhidos por dois microfones. Consideraram-se, ainda, duas situações distintas: fonte sonora conhecida *a priori* e desconhecida. Na primeira situação utilizaram-se *matched filters* enquanto na segunda se efectuou a correlação cruzada. Ambos os métodos foram testados com êxito num pequeno robot móvel e autónomo. Conforme seria de esperar, a utilização de uma fonte sonora pré-estabelecida em conjugação com a técnica de *matched filtering* e a SPLL mostrou resultados superiores, tendo o robot sido capaz de seguir a fonte sonora mesmo com níveis de intensidade sonora do farol praticamente no limiar da audição humana. A técnica de correlação cruzada também funcionou mas existiram situações em que o robot se perdia, e.g. em ambientes com muita reverberação. Nesta situação, e também conforme seria de esperar, os resultados eram melhores com sons abruptos e secos, tais como palmas, batidas com um objecto duro numa mesa, etc.

Referências

- [1] Jie Huang, et al., *A Model-based Sound Localization System and its Application to Robot Navigation*, Robotics and Autonomus Systems, Vol.27, pp.199-209, 1999.
- [2] William M. Hartmann, "How We Localize Sound", *Physics Today*, pp.24-29, Nov., 1999.
- [3] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, 1984.
- [4] R. E. Best, *Phase-Locked Loops*, 3ª ed, McGraw-Hill, 1997.
- [5] Chiang-Jung Pu, *A Neuromorphic Microphone for Sound Localization*, PhD Thesis, University of Florida, 1998, USA.