

二次元 CSP による目領域探索の高速化 A Fast Algorithm for Eye Detection Using Two-Dimensional CSP

鈴木 亜紀子[‡]
Akiko SUZUKI

滝口 哲也[†]
Tetsuya TAKIGUCHI

有木 康雄[†]
Yasuo ARIKI

1. はじめに

人間の顔，特に目は，注意や意図，感情といった心理状態に応じて様々な変化を表出するため，その変化を認識することは有用である．これまで，一次元の CSP (Crosspower-Spectrum Phase) は 2 つのマイクロフォン間に到達する信号の時間差を推定する方法として提案され [1] [2][3]，雑音下において有効な手法であると報告されている．本稿では，画像のテンプレートマッチングに適用した新しい二次元 CSP 法を提案，カメラ映像を用いることで非拘束的に，目領域を高精度でかつ高速に抽出することを目指す．

2. 正規化相互相関

正規化相互相関法では，全探索領域内に対してテンプレート画像と相関を求め，最も相関の高かった領域（位置）を対象物領域と見なす．画像サイズ $W \times H$ の探索対象画像 $I(x, y)$ と， $w \times h$ のテンプレート画像 $T(i, j)$ に対して，相関係数 $R(x, y)$ を計算する．

$$R(x, y) = \frac{\sum_{i,j} \{T(i, j) - \bar{T}\} \cdot \{I(x+i, y+j) - \bar{I}\}}{\sqrt{\sum_{i,j} \{T(i, j) - \bar{T}\}^2 \sum_{i,j} \{I(x+i, y+j) - \bar{I}\}^2}} \quad (1)$$

ここで，

$$\bar{T} = \frac{\sum_{i,j} T(i, j)}{w \cdot h} \quad (2)$$

$$\bar{I} = \frac{\sum_{i,j} I(x+i, y+j)}{w \cdot h} \quad (3)$$

である．これにより，マッチングした領域の先頭座標 (\hat{x}, \hat{y}) は

$$(\hat{x}, \hat{y}) = \underset{x,y}{\operatorname{argmax}} R(x, y) \quad (4)$$

となる．

3. 二次元 CSP による物体探索

図 1 に二次元 CSP 法の処理の流れを示す．まず，探索対象画像 $I(x, y)$ とテンプレート画像 $T(i, j)$ に対して，フーリエ変換を適用し周波数領域へ変換する．

$$I(\omega_1, \omega_2) = \sum_{x,y} I(x, y) e^{-j\omega_1 x} e^{-j\omega_2 y} \quad (5)$$

$$T(\omega_1, \omega_2) = \sum_{i,j} T(i, j) e^{-j\omega_1 i} e^{-j\omega_2 j} \quad (6)$$

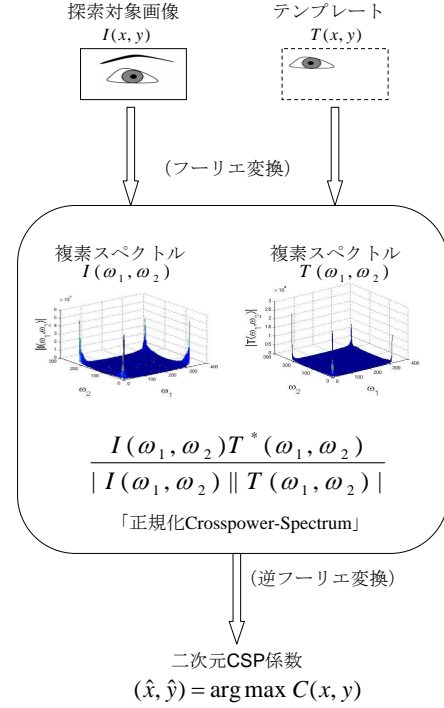


図 1: 二次元 CSP の求め方

ただし，各々の画像サイズが異なるので，小さい画像サイズにゼロパディングを行い，同じサイズにしてから二次元フーリエ変換を適用する．

得られた各々の複素スペクトルを用いて，正規化 Crosspower-Spectrum は以下の式で求められる．

$$\phi(\omega_1, \omega_2) = \frac{I(\omega_1, \omega_2) T^*(\omega_1, \omega_2)}{|I(\omega_1, \omega_2)| |T(\omega_1, \omega_2)|} \quad (7)$$

となり，最後に式 (7) に逆フーリエ変換を適用し，二次元 CSP 係数を求める．

$$C(x, y) = \sum_{\omega_1, \omega_2} \phi(\omega_1, \omega_2) e^{j\omega_1 x} e^{j\omega_2 y} \quad (8)$$

(式 (8) において定係数は無視した.) よって，マッチングした領域の先頭座標 (\hat{x}, \hat{y}) は

$$(\hat{x}, \hat{y}) = \underset{x,y}{\operatorname{argmax}} C(x, y) \quad (9)$$

となる．

実験で得られた二次元 CSP の値を図 2，正規化相互相関係数の値を図 3 に示す．二次元 CSP では，周波数上にて画像を白色化することにより，位相情報のみに

[†]神戸大学工学部

[‡]神戸大学大学院自然科学研究科

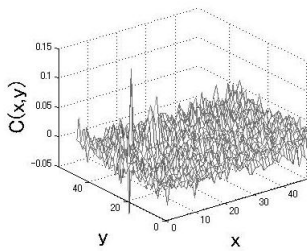


図 2: 二次元 CSP 係数 .

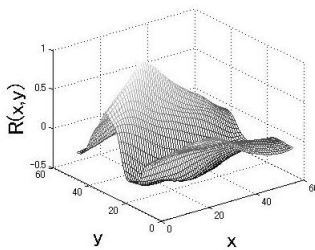


図 3: 正規化相互相関係数 .

注目した探索が可能となり，従来の信号領域における正規化相互相関法よりも，目的とする領域（位置）に対して鋭いピーク値を得る事が出来ているのが分かる．

4. 両目の抽出実験

4.1 実験条件

室内環境にて 1 人の男性話者に対して，デジタルビデオカメラ（Victor GR-HD1）により収録を行った．

映像データに対して，まず AdaBoost[4] を使い顔の切り出しを行った（顔の切り出しには OpenCV に含まれる検出器を使用した）．次に目領域の探索は，切り出された四角領域内の上部分の領域内に対して行う．目の探索領域を 2 等分にして，右側を右目のテンプレートを使い探索を行い，左側を左目のテンプレートを使い目領域の探索を行った．評価データの総フレーム数は，3137 フレームである．

4.2 テンプレート作成

テンプレートを作成するために，まず Haar 型特徴量 [4] を使用し，AdaBoost を用いて目領域の抽出を行う．今回，Haar 型学習に使用した目領域データ数は 3000 枚，非目領域データは 1732 枚である（評価話者，及び収録環境は異なる）．ただし，この検出器を用いただけでは，誤検出も多く含まれてしまうため，評価データに対して最初の 100 フレーム分の目領域を抽出し，それぞれの目領域を 3 分割し，中央領域とその両側の矩形内平均輝度差が最も大きいものを，正しい目領域であると考えテンプレートとして採用した．抽出されたテンプレートを図 4 に示す．

4.3 実験結果

探索結果（正解率）を表 1 に示す．正規化相互相関法（NCC: Normalized Cross Correlation）で左目，右

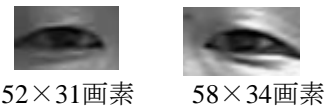


図 4: 自動生成されたテンプレート .

表 1: 従来の信号領域における正規化相互相関法（NCC:Normalized Cross Correlation）と，二次元 CSP（2D-CSP）の実験結果 .

	左目精度 (%)	右目精度 (%)	処理時間 (sec)
NCC	100.0	100.0	1372
2D-CSP	98.6	97.7	323

目ともに 100%，一方，二次元 CSP では左目 98.6%，右目 97.7% となった．テンプレートが開眼状態であるため，二次元 CSP では目が閉まる際の誤検出が存在した．

評価データの処理時間（Pentium 4, CPU 3GHz）は，正規化相互相関法（NCC）で約 1372 秒であり，一方，二次元 CSP では約 323 秒となった．計算時間は 1/4 程度まで削減されている．従来法では，画素単位で探索信号領域を移動しながら相関を計算していたのに対し，二次元 CSP では周波数上にて相関処理を行うため，計算量が大幅に削減され高速な探索が可能となっている．また，探索目領域の画素数は平均約 8900 画素であった．

5. おわりに

本稿では，二次元 CSP による物体領域の抽出手法を検討した．二次元 CSP では，周波数上にて画像を白色化することにより，位相情報のみに注目した探索が可能となり，従来の信号領域における相互相関法よりも，目的とする領域（位置）に対して鋭いピーク値を得る事が出来る．また，周波数上にて相関計算を行うため，高速な探索が可能となる．今後は，マルチテンプレート化，帯域分割化など，探索精度の改良を進めていく．

参考文献

- [1] C. H. Knapp, G. C. Carter, "The Generalized Correlation Method for Estimation of Time Delay," IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-24, pp. 320-327, Aug. 1976.
- [2] M. Omologo, P. Svaizer, "Acoustic Event Localization Using a Crosspower-Spectrum Phase Based Technique," Proc. ICASSP, vol. 2, pp. 273-276, 1994.
- [3] 西浦敬信, 山田武志, 中村哲, 鹿野清宏, "マイクロホンアレーを用いた CSP 法に基づく複数音源位置推定", 電子情報通信学会論文誌, Vol.J83-D-II, No.8, pp. 1713-1721, Aug. 2000 .
- [4] P. Viola, M. Jones, "Robust Real-Time Face Detection," International Journal of Computer Vision 57 (2), pp. 137-154, 2004.