

416 独立成分解析を用いた混合画像の分離について

About Separation of a Mixture of Independent Images by Independent Component Analysis

正 松浦 勉 (群馬大)

Tsutomu MATSUURA, Gunma University, 1-5-1 Tenjin-cho Kiryu Gunma

Keywords: ICA, blind source separation, Kullback-Leibler information, neural network

1. 緒言

我々は、人々の話し声やさまざまな音が重なり合ったパーティー会場の中にあっても、そうした雑音に煩わされることなく特定の人の話し声を聞き分け、理解することができる。このように重なり合った信号の中から特定の信号を識別する能力は“カクテルパーティー効果”と呼ばれ、Blind Source Separationの典型的な例でありさまざまな研究がなされている。独立成分解析(Independent Component Analysis; ICA)とは、Blind Source Separationへの応用、生体の初期視覚における自己組織化のモデル化、脳波による脳活動解析への応用などが研究されている新しいデータ解析手法のことである。

ICAの最大の特徴は、複数の原信号間の独立性のみを仮定し、それ以外の情報はすべて未知であるということである。また、信号と同じ数、またはそれ以上の異なった場所での観測が同時に行えることも仮定する。

本報告では、互いに独立な原信号(元画像)をある割合で混合した混合画像に対し独立成分解析を行うことによって、再び元の画像を復元できるかということを、単純なサンプル画像と、応用として実際の写真とを用いて解析を行う。また、その学習アルゴリズムの有効性、安定性を検証し考察する。

2. 独立成分解析とは

ICAは主成分分析(PCA)と同様、多次元信号の解析手法である。PCAでは信号の2次の相関のみに注目し、それら無相関にする変換を求め、これに対しICAは高次の統計量、あるいは時間的な相関に基づく独立性により信号を分離する変換を求め、正規分布以外の確率分布に対しては一般に無相関と独立性とは一致しない。したがってPCAとICAは異なる結果を与える。

観測値及び信号源をベクトル $\mathbf{x}(t)=(x_1(t), \dots, x_m(t))^T$, $\mathbf{s}(t)=(s_1(t), \dots, s_n(t))^T$ で表現すると、 $\mathbf{x}(t)=\mathbf{A}\mathbf{s}(t)$ と書かれる(\mathbf{A} : $m \times n$ の混合行列)。ここで以下の仮定をおく。1) 信号源 $\mathbf{s}(t)$ の成分は互いに独立。2) 混合行列 \mathbf{A} は正則。3) 観測の数 m と信号源の数 n は等しい。

ICAは上記諸量のうち、 \mathbf{A} , \mathbf{s} を両者ともに(上記の仮定が成り立つこと以外に)全く知らない状況下において、 $\mathbf{s}(t)$ の独立性を頼りにして観測値 $\mathbf{x}(t)$ から $\mathbf{s}(t)$ を推定する(Fig.1の \mathbf{y})方法を与える。

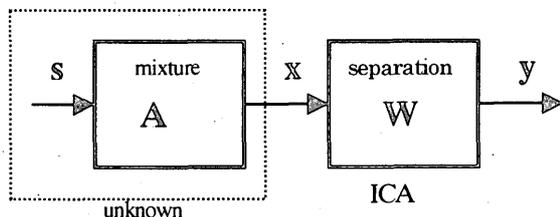


Fig.1 Model of ICA(mixture and separation)

信号源を復元するため、分離モデルとしての単層の線形フィードフォワード神経回路網: $\mathbf{y}(t)=\mathbf{W}(t)\mathbf{x}(t)$, あるいは、リカレント神経回路網: $\mathbf{y}(t)=\mathbf{x}(t)-\mathbf{V}(t)\mathbf{y}(t)$ を導入する。ここで、出力ベクトルは $\mathbf{y}(t)=(y_1(t), \dots, y_n(t))^T$ である。 $\mathbf{W}(t)$ と $\mathbf{V}(t)$ はそれぞれの神経回路網の結合荷重で、 $\mathbf{W}(t)=(\mathbf{I}+\mathbf{V}(t))^{-1}$ ならば(\mathbf{I} は単位行列)、両神経回路網は同一なものと思なすことができる。

上式から、神経回路網の出力 $\mathbf{y}(t)$ と信号源 $\mathbf{s}(t)$ の関係は

$$\mathbf{y}(t)=\mathbf{W}(t)\mathbf{A}\mathbf{s}(t)=\mathbf{G}(t)\mathbf{s}(t) \quad (1)$$

となる。ここで $\mathbf{G}(t)$ は信号源から出力への合成行列である。仮定より $\mathbf{s}(t)$ の成分が独立で、 \mathbf{A} の逆行列が存在していることから、出力 $\mathbf{y}(t)$ の成分が独立となるのは、上式の合成行列が、単位行列と等しくなるか、あるいは、 $\mathbf{G}=\mathbf{D}\mathbf{P}$ となるかの二つの場合が考えられる。ここで対角行列 $\mathbf{D}=\text{diag}(d_1, \dots, d_n)$ の要素は非ゼロであり、 \mathbf{P} は置換行列である。(各行、各列に1が1個ずつありその他の成分は総てゼロ。)

前者の場合は原信号が(順番、大きさも含めて)忠実に再現されたことになる。後者の場合は出力の各成分は一般に $y_i(t)=d_i s_j(t)$ となり、分離(復元)された信号は原信号の順序とスケールを保存しない(原信号の独立性のみを仮定し、出力信号の独立性のみを要請している)のであるから当然の帰結である。)

以上より混合行列 \mathbf{A} が既知ならば、 $\mathbf{W}=\mathbf{A}^{-1}$ と置けば原信号を簡単に分離(復元)できるが、我々は \mathbf{A} を知らないで、いかにして $\mathbf{W}(t)$ を \mathbf{A}^{-1} あるいは $\mathbf{D}\mathbf{P}\mathbf{A}^{-1}$ に近づけるかという学習方法を探り出す問題となる。

3. Kullback-Leibler 情報量最小化に基づく学習アルゴリズム

n 変数 $\mathbf{y}=(y_1, \dots, y_n)^T$ の同時分布、周辺分布間の Kullback-Leibler 情報量(K-L情報量)は

$$I(\mathbf{y})=I(y_1, \dots, y_n)=\sum_{i=1}^n H(y_i)-H(\mathbf{y}) \quad (2)$$

結合分布のエントロピー $H(\mathbf{y})$ は $\mathbf{y}=\mathbf{W}\mathbf{x}$ より $P_{\mathbf{y}}(\mathbf{y})=P_{\mathbf{x}}(\mathbf{W}^{-1}\mathbf{y})\det(\mathbf{W})$

と書けることを用いると、

$$H(\mathbf{y})=-\int P_{\mathbf{y}}(\mathbf{y})\log P_{\mathbf{y}}(\mathbf{y})d\mathbf{y}=H(\mathbf{x})+\log|\mathbf{W}| \quad (3)$$

となる。ここで \mathbf{x} と \mathbf{y} が1対1で対応しているので \mathbf{x} で平均を取ることと \mathbf{y} で平均を取るとは同値であることに注意する。よって(2),(3)より K-L情報量は

$$I(\mathbf{y})=\sum_{i=1}^n H(y_i)-H(\mathbf{x})-\log|\mathbf{W}| \quad (4)$$

となる。 \mathbf{y} の各要素が互いに独立になるような変換 \mathbf{W} を求めたいわけだが、このとき \mathbf{y} の同時分布は各要素の周辺分布の積に一致し、同時に K-L情報量は最小になる。よって求めたい \mathbf{W} は $\partial(\mathbf{y})\partial\mathbf{W}=\mathbf{0}$ を満たす。このような \mathbf{W} を求めるために、(4)を \mathbf{W} で偏微分すると、(\mathbf{x} のエントロピーは \mathbf{W} とは独立であることに注意)第3項は $\partial\log|\mathbf{W}|/\partial\mathbf{W}=(\mathbf{W}^{-1})^T$ となり、各 y_i のエントロピーは

$$\begin{aligned} \frac{\partial H(y_i)}{\partial \mathbf{W}} &= \frac{\partial H(y_i)}{\partial w_j} = -\frac{\partial E_{P_{\mathbf{y}}}[\log P_{\mathbf{y}}(y_i)]}{\partial w_j} = -\frac{\partial E_{P_{\mathbf{y}}}[\log P_{\mathbf{y}}(\sum w_a x_a)]}{\partial w_j} \\ &= -E_{P_{\mathbf{x}}}\left[\frac{d \log P_{\mathbf{y}}(y_i)}{dy_i} x_i\right] = E_{P_{\mathbf{x}}}(\phi(\mathbf{y})\mathbf{x}^T) \end{aligned} \quad (5)$$

ただし、(5)で第2項から第5項は行列(最左辺)の要素で表示しており、最右辺の ϕ の各要素は第5項の括弧内で定義されるものである。これらをもとに $\partial(\mathbf{y})\partial\mathbf{W}=\mathbf{0}$ を \mathbf{W} に関して解けば復元のための \mathbf{W} が得られるはずだが、 ϕ の非線形性のために解析的に解くことはできず、反復法等で数値的に求めざるを得ない。最も単純な方法は最急降下法である。 ε を小さな数として \mathbf{W} を $\mathbf{W}+\varepsilon\mathbf{U}$ と変

化させたときのUのノルムを

$$\|U\|_w^2 = \|U\|_w^2 = \text{tr}(W^{-T}U^T U W^{-1}) = \text{tr}(W^{-1}W^{-T}U^T U) \quad (6)$$

で定義し、K-L情報量を減らす勾配を求め最急降下法を用いると、

$$\Delta W = (I - E_{p_i} [\phi(y_i)y_i^T])W \quad (7)$$

となるが、実際はこれをオンライン化した(8)を用いる。

$$\Delta W_{i+1} = \eta_i (I - \phi(y_i)y_i^T)W_i \quad (8)$$

4. シミュレーション結果と考察

原画像として Fig.2 の 200×200 の画素から成り立っている 3 枚の画像を用いた。3 枚目はランダム画像である。これらを混合ベクトル(0.3,0.1,0.9), (0.3,0.1,0.9), (0.3,0.3,1.0)で混合したものが Fig.3 である。(これらは線形独立に選んである。)我々に与えられているのは Fig.2 の 3 枚の画像と、「それらは未知の(独立な)原画像が何らかの線形混合によって得られたものである。」という知識だけである。ICA を実施するに際してデータの前処理による結果の違いを見るために、p1) データをそのまま使う(前処理をしない。), p2) データを標準偏差で正規化する, p3) 最大値で正規化する, p4) データを標準偏差で正規化した後、最大値で割る、の 4 通りの前処理を行なった。

Fig.4 は(8)の $\eta=0.0003$, 前処理 p3 で ICA を行なった結果である。分離結果は(出力の順序は交換しているが)殆ど原画像と見まがうほどである。この場合に学習で得られた W が DPA^{-1} の形になっているかを確認するために、 $W A$ を計算してみると(ただし A は前処理も含めたので前述した混合ベクトルを変形したものになっている。), $W A =$

$$\begin{bmatrix} 0.3339 & -0.3619 & -0.0319 \\ 0.1226 & 0.0841 & 0.0900 \\ 0.8779 & -0.0146 & 1.0 \end{bmatrix} \begin{bmatrix} -2.115 \cdot 10^{-3} & -7.050 \cdot 10^{-4} & -6.1345 \cdot 10^{-3} \\ 2.101 \cdot 10^{-3} & -7.100 \cdot 10^{-4} & -6.303 \cdot 10^{-3} \\ 1.1646 \cdot 10^{-3} & 1.646 \cdot 10^{-3} & 5.1485 \cdot 10^{-3} \end{bmatrix} \\ = \begin{bmatrix} -0.0015 & -3.440 \cdot 10^{-5} & -1.224 \cdot 10^{-5} \\ 6.565 \cdot 10^{-5} & 2.862 \cdot 10^{-5} & -0.0008 \\ -0.0002 & 0.001 & 7.372 \cdot 10^{-5} \end{bmatrix} \cong \text{diag} \begin{bmatrix} 0.0015 & & \\ & -0.0008 & \\ & & 0.001 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

となっており W が DPA^{-1} の形になっていることが確認される。

前処理については p1 の場合は学習の収束が遅く、なおかつ収束も遅い傾向がみられたが、他の 3 つの場合は収束、分離結果とも良好であり、本方法の有効性が確認できた。

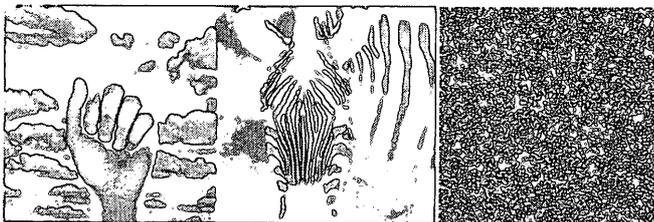


Fig.2 Original images (unknown)

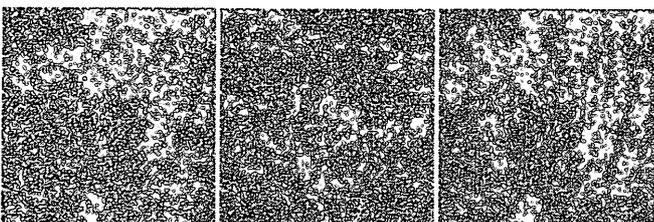


Fig.3 Mixed images

次に、混合画像の一部を用いて元画像の再生を試みた。これはそれぞれの元画像についてはその総ての部分に対して同一の割合で混合して混合画像を得ているのであるから、与えられた混合画像の一部だけからでも、その混合比情報が得られ筈だと考えられるからである。Fig.5 は 3 枚の混合画像(200×200)の 31 行目から 80 行目までの部分を取りだし、それについて ICA を行ない、学習によって得られた分離行列を混合画像全体に適用して得られた分離画像である。分離された「手の画像」にかすかに「シマウマ」のシマが重なっているのが見て取れる。そこで元画像のうち 2 枚の「意味のある画像」(手とシマウマ)について「独立性検定」を行なったところ、たまたまその部分(31 行~80 行)は「独立性」が強くないことがわかった。そこで「連続な 50 行分を切り出す」という条件の元に「独立性の強い部分」を探してみると、74~123 行の部分と 81~130 行の部分であった。それらの部分からの ICA によって上述のように分離画像を構成してみると、むしろ混合画像全体に対し ICA を施した場合より収束も早く(約 8 倍)分離の程度も良好であった。これは元画像を局所的に見たときに偶然に「独立性の弱い部分」が混在していることが考えられ、その部分が学習の収束を阻害していると考えられる。これらの詳細なメカニズムや、「どこまで小さい部分から元画像が再生か?」については現在検討中である。

5. 結言

- 1) この論文で提示した学習アルゴリズムを用いてシミュレーションを行った結果、混合画像を元の 3 枚の画像に復元できることを確認し本方法の有効性が示された。
- 2) 復元画像の分離の良し悪しは、学習係数 η に依存していることを確かめた。しかし、最適な η を見つけるためには繰り返しシミュレーションを行う必要がある。
- 3) データの前処理として 4 通りの方法(p1~p4)を採用したが、p1 以外はどの方法による結果も大きな違いはなかった。
- 4) 復元画像が良く分離されている場合は、学習によって得られた重み W は、 DPA^{-1} を十分良く近似していることを確認した。
- 5) 混合画像の一部を用いた解析による復元が可能であることを確かめた。このことから大画面を扱う際には、その一部を用いた解析を行うことによって計算時間の大幅な短縮が期待できる。

参考文献

- 1) S. Amari, T. Chen, A. Cichocki, Stability analysis of adaptive blind source separation, Neural Networks, 10-8, pp1345-1351 (1997)
- 2) Te-Won Lee, Independent Component Analysis, Kluwer Academic Publishers, Boston (1998)

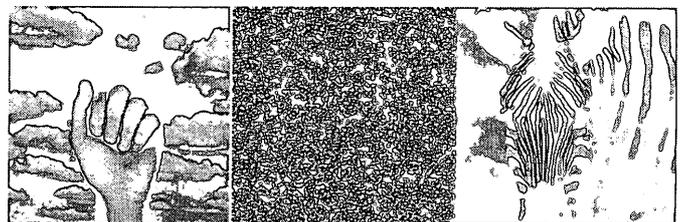


Fig.4 Result of separation by ICA

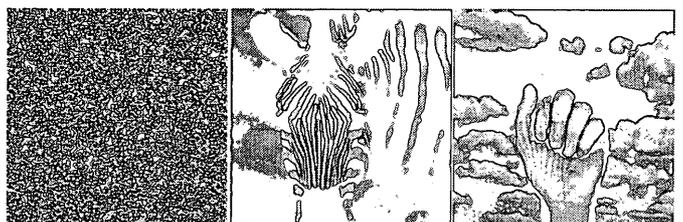


Fig.5 Result of separation by ICA (using partial images)