

ガウシングラフィカルモデルを用いた動画補修の提案 Inpainting of moving images based on Gaussian graphical model

藤田 涼世†
Ryosei Fujita

大久保 潤†
Jun Ohkubo

1. はじめに

本研究では、高速に移動している際にカメラについた水滴を除去することを目的とした手法を提案する。動画補修は様々な手法が提案されているが、計算時間がかかるため、リアルタイムの補修が困難である[1]。そこで今回はリアルタイムに処理をすることを念頭におき、計算量が少なく済む手法を組み合わせることで手法を構成する。水滴箇所を検出については2つの線形分離機を組み合わせることにより、高速な検出を実現する。また、ガウシングラフィカルモデルを用いた静止画画像補修の手法を動画に拡張して、検出した水滴箇所の補修をおこなう。

2. 水滴箇所の判別

水滴箇所は色や形が常に一定ではないという性質がある。そのため、特徴量を用いた判別が難しい。そこで、今回は動画中の静止画を細かく区切り、離散コサイン変換をおこなって、その時間差分を用いることにより、水滴箇所の判別をおこなった。

2.1 離散コサイン変換

離散コサイン変換は離散信号を周波数領域に変換する方法の一種であり、画像に用いることで、対象の画像を周波数成分で表すことができる。水滴箇所は建物などの背景と比べ高周波成分が少ないため、離散コサイン変換を行うことにより一定の判別能力を見込むことができる。しかし空などのカメラの移動中でも代わり映えのしない背景箇所も高周波成分は少なくなるため、水滴との判別が困難となる。そこで今回は、離散コサイン変換をおこなった画像と1フレーム前静止画に対して離散コサイン変換をおこなった画像との差分を求めることにより、水滴箇所の判別をおこなった。水滴はカメラが高速移動している限りその形が変形するため、低周波成分の値はフレームごとに異なる。そのため前フレームとの差分を求めると低周波成分の値にわずかな誤差がでる。対して空などの低周波のみで構成される背景は前フレームとの差分を求めても差がほとんど生まれない。

2.2 線形分離機

線形分離機では入力 $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$ に対して重み $\mathbf{w} = \{w_1, w_2, \dots, w_n\}$ を掛け合わせたものを活性化関数 f に入れることで、入力の判別をおこなうことができる。今回の例では離散コサイン変換した静止画の時間差分を入力とし、出力は 0, 1 の二値分類とする。図1に線形分離機のモデルを示す。このモデルの場合、出力 z は以下の式で求められる。

$$u = x_1 w_1 + x_2 w_2 + \dots + x_n w_n + b$$

$$z = f(u)$$

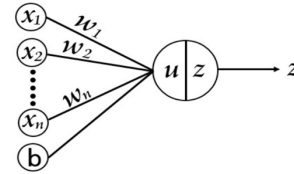


図1:線形分離機のモデル

この時の b はバイアスを表す。活性化関数 f は通常は単調増加する非線形関数を使用される。今回は活性化関数にロジスティック関数を用いた。式は以下のように表される。

$$f(u) = \frac{1}{1 + e^{-u}}$$

線形分離機の重み w は教師あり学習で求める。通常線形分離機の層を深くすることで検出の精度をあげることが可能となるが、今回は計算量削減のため線形分離機を以下の図2のように組み合わせることにより水滴の検出率を高めた。

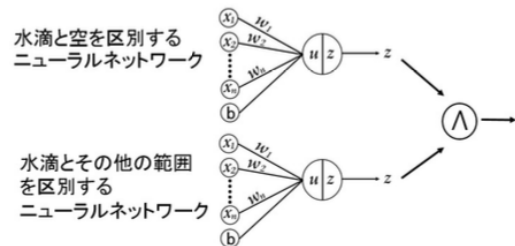


図2:線形分離機を二つ組み合わせた図

一つ目の線形分離機では水滴箇所と空などの代わり映えのしない背景箇所の判別をおこない、二つ目の線形分離機では水滴箇所と建物などがある背景箇所の判別をおこなえるよう学習させた。どちらの線形分離機も水滴と判別をした時のみその範囲を水滴と判断する。

3. 水滴箇所の補修

3.1 静止画補修

水滴箇所の補修にはガウシングラフィカルモデルを用いた静止画補修の方法を用いる[2]。この方法は欠損してしまった範囲のピクセルの値を隣接するピクセルの値を用いて予測する方法である。説明のために以下にモデルの図を記載する。

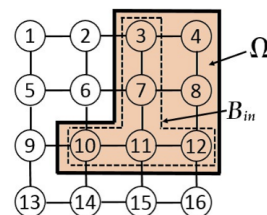


図3:画像のピクセルのモデル

† 埼玉大学, Saitama University

図3は画像のピクセルを表す。太枠で囲まれた範囲 Ω が欠損範囲とする。また点線で囲まれている、 Ω 内の欠損していないピクセルデータと隣あった範囲を B_{in} とし、 B_{in} と隣り合う、欠損していない範囲のピクセルを N_D で表す。各ピクセルの値は上下左右隣り合ったピクセルからのみ互いに影響し合っていること、また近くのピクセル値はお互いに近い値を持ちやすいという仮定をし、コスト関数として以下の式を与える。

$$E(f) \equiv +\frac{1}{2} \sum_{i \in B_{in}} \beta_i (f_i - z_i)^2 + \frac{1}{2} \sum_{(ij) \in N_D} \alpha_{ij} (f_i - f_j)^2$$

$$z_i \equiv \frac{1}{|N_D(i)|} \sum_{j \in N_D(i)} g_j$$

各記号の意味は以下の通り。

- i, j ... ピクセルの番号
- g_j ... j 番のピクセルの値
- β_i ... パラメータ定数
- α_{ij} ... パラメータ定数
- f_i ... Ω 内のピクセルの値

コスト関数の右辺の左側の項は、 B_{in} 内のピクセルが受ける欠損していない範囲からの影響力を表しており、右側の項は、 Ω 内にある隣り合ったピクセル同士の影響力を表している。この影響力は β 、 α を変更することによって調整可能である。この最適化問題は通常全ての値の組み合わせから探すとなると非常に大量な計算が必要とされ、一般に計算困難になる。そこで文献[2]では確率モデルの定式化を行うことにより、近似法を使用して少ない計算時間でコスト関数を最小値に近づけることが検討された。確率モデルにおけるコスト関数の定式化は以下の通りである。

$$P(f) = \frac{1}{Z} e^{-E(f)}$$

Z は規格化定数を表す。この確率を最大にする f が、コスト関数を最小にする f と一致することになる。静止画の各ピクセル値を整数ではなく実数とみなすことにより、確率モデルを多次元ガウス分布に帰着することができ、以下の式を得ることができる。

$$P(f) = (2\pi)^{-\frac{|N|}{2}} (\det A)^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (f - m)^T A (f - m) \right\}$$

ここで、各変数は次の式で定義される。

$$A_{ij} \equiv \begin{cases} -\alpha_{ij} & ((ij) \in N_D) \\ \tilde{\beta} + \sum_{k \in N_D(i)} \alpha_{ik} & (i = j) \\ 0 & (\text{otherwise}) \end{cases}$$

$$m \equiv A^{-1} b$$

$$\tilde{\beta} \equiv \begin{cases} \beta_i & (i \in B_{in}) \\ 0 & (i \notin B_{in}) \end{cases}$$

$$b_i \equiv \tilde{\beta}_i z_i$$

確率モデルを最大とする f を選ぶには、 $f = m$ とすれば良いのは明白である。ベクトル m を計算する際、行列の逆行列を求めるのは計算時間がかかる。しかし、問題を連立方程式の形に定式化できることを利用すると、ヤコビ法、

またはガウス・ザイデル法を用いて高速に解くことができる。例えば以下の式の形である。

$$m_i = \frac{1}{A_{ii}} (\tilde{\beta}_i z_i + \sum_{j \in N_D(i)} a_{ij} m_j)$$

こうして得られた m の値を一番近い整数に変換することで、画像の補修が完成する。

3.2 動画補修

動画を補修する際は、補修したいフレームの一つ前のフレームの情報も使い補修を行う。1フレームごとの物体の動く距離を測定し、前フレームの情報を測定した距離だけ動かし、補修を行う。補修は図4のイメージのように、ピクセル間のつながりを、前フレームと三次元的に繋がっていると想定することによって補修をおこなう。

具体的な設計としては以下の式を用いる。

$$z_i \equiv \frac{1}{|N_D(i)| + 2} (g_i^t + g_k^{t-1} + \sum_{j \in N_D(i)} g_j^t)$$

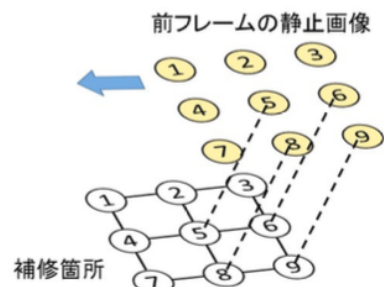


図4:前フレームを利用した画像補修のイメージ

g_i^t は t フレームでのピクセルの値を表し、 k は現フレームで i に対応する前フレームのピクセルの値を表す。また、前フレームのデータだけに依存するのではなく、水滴下の情報も取り入れるために g_k^t の値も z に加えている。

4. おわりに

本研究では線形分離機を組み合わせることで水滴の検出をおこなえること、ガウシアングラフィカルモデルを用いることで動画の補修をおこなう枠組みを提案した。動画の補修は速い補修を可能とする枠組みであり、計算回路をFPGAなどで専用のものとすることでさらに速い動画補修をおこなえると考えられる。

謝辞

本研究で使用したデータの一部は、カルソニックカンセイ株式会社からご提供いただきました。

参考文献

[1] K.A.Patwardhan, G.Sapiro, and M.Bertalmio: Video Inpainting Under Constrained Camera Motion, proceedings of IEEE Transactions on Image Processing, vol. 16, no. 2:545-553(2007)
 [2] M.Yasuda, J.Ohkubo and K.Tanaka: Digital Image Inpainting based on Markov Random Field, Proceedings of the International Computational on Computational Intelligence for Modelling, Control and Automation and International Conference Intelligent Agents, Web Technologies and Internet Commerce,2:747-752(2005).