

時空間画像解析を利用した 車載カメラ映像からの障害物除去手法

栗林 宏 輔^{†1} 小野 晋太郎^{†2}
川崎 洋^{†1} 池内 克史^{†2}

都市などで車載したカメラで撮影した映像を街のモデリングなどに利用するシステムが多く提案されるようになってきている。そのような映像には歩行者や電柱、街路樹などモデリングの際に障害物となる物体が多数写りこんでいる。そこで、本論文では時空間画像解析を利用して、そのような障害物を効率的に除去する手法を提案する。特に、本論文では、撮影対象物としてビルや家屋など、市街地が主に道路に平行した平面構造で構成されることに着目し、時空間画像解析である EPI 解析を用いて、障害物を自動的に除去する手法を提案する。時空間解析を利用することで、従来手法では難しかった屋外環境における歩行者や街路樹、電柱など複雑な障害物を効率よく除去することができる。実際に屋外で実験を行ったところ、正しく障害物を除去することができた。

Efficient removal of obstacles from car-mounted video data by using spatio-temporal analysis

KOUSUKE KURIBAYASHI,^{†1} SHINTARO ONO,^{†2}
HIROSHI KAWASAKI^{†1} and KATSUSHI IKEUCHI^{†2}

In this research, the method which can automatically remove objects, such as pedestrians, telegraph poles, roadside trees, etc., from the on-vehicle video camera is proposed; such images are widely used for urban scene modeling (e.g. Google Street View), and a removal of these objects now becomes a critical issue. Since an input data is a video stream and an urban scene is mainly composed of planar surfaces parallel to the street, the method can effectively remove the objects by using the spatiotemporal image analysis. To show the strength of the method, several experiments using real data are conducted, which resulted in a successful removal of complicated objects.

1. はじめに

街などの広域空間をモデリングするために、ビルなどの建造物のみを撮影しようとしても、その建物の前にある樹木や電柱などが映り込み、対象とする建造物だけの画像が手に入らないことがよくある。このように撮影した画像上に対象物と障害物が同時に重なりあって映りこむ状況は日常的に見られる現象である。この現象は撮影者の意図通りの画像取得を妨げるだけでなく、コンピュータビジョンにおけるトラッキングや物体計測などの処理や、コンピュータグラフィックスにおける3次元都市構築における実世界のテクスチャ取得などの際にも大きな問題となる。また、最近では Google Street View などのシステムにおいて映り込んでしまった人や自動車などは、プライバシーの問題なども引き起こしている。

そこで、本研究は、撮影対象物と障害物が重なったシーンから障害物を除去し、撮影対象物のみを映し出すことを目的とする。この問題の解決策として、時空間画像解析の一つである EPI 解析を用いる。EPI 上では現実世界のシーンにおける3次元点が直線となって現れるため、これを利用すれば、障害物が複雑な形状や移動物体や複数個であっても、直線上の処理により障害物除去を容易に行うことができる。また、EPI は等速直線運動を仮定しているが、市街などで撮影する際にはこの仮定が成り立ちにくいいため、カメラが等速運動ではない場合において、これを自動補正する手法もあわせて提案する。

実際に車載カメラで市街のデータを撮影し実験を行ったところ、電柱や電線など、建物を撮影するにあたり障害物と考えられる対象を自動で除去することが出来た。以下では、関連研究について述べ、次に提案手法の詳細を紹介し、最後に、実験および結論を述べる。

2. 関連研究

障害物に限らず、画像や映像上の特定の領域を違和感無く穴埋めすることは、需要も大きく長く研究されてきた。最近では、グラフィックス分野において、lazy snapping による簡易なインタラクティブ操作により、周辺のデータを用いて障害物領域を違和感無く埋める手法 (inpainting) などが発表されている³⁾。また、さらにこれを拡張し、ビデオ映像から不要な部分を除去する手法も提案されている⁶⁾。以下では、このような過去の障害物除去に関する研究を、空間および時間的な補間によるものに分類して紹介する。

^{†1} 埼玉大学 Saitama University

^{†2} 東京大学 University of Tokyo

2.1 空間的除去手法

単画像において、ある特定の領域を周囲の情報で穴埋めする手法は、空間的な補間による障害物除去と考えることができる。小幡らは、単一画像から画像の局所的な類似性とフラクタル性を利用し、不要オブジェクトを囲む僅かなベクトルと類似する領域を探索して推定、修復した¹⁾。榎本らは、複数視点から対象物を撮影し、その画像を射影変換することで比較し、障害物を除去した²⁾。Liらは、lazy snapping による簡易なインタラクティブ操作により選択した領域をグラフィカットにより高精度に推定し、周辺のデータを用いて滑らかに補間する手法を提案している³⁾。

2.2 時間的除去手法

固定カメラによる背景作成の研究などは、時間的な補間による障害物除去と考えることができる。このような研究はこれまで非常に多く行われている。例えば、Korahらはモーションキューを使用し周りから補間する方法を提案した⁴⁾。また、Matsushitaらは、屋外に固定カメラで撮影した交差点映像から光源(太陽)の動きを考慮することで、車両などの障害物を効率よく除去する手法を提案している⁵⁾。

2.3 時空間を考慮した除去手法

ビデオ映像などを対象として、時空間的な補間により障害物を除去する手法が最近提案されるようになってきている。Shiratoriらは、ビデオ映像においてグラフィカットを用いて、特定の対象物体を除去し時空間内で類似のセグメントを探索することで、滑らかにビデオ映像を補間する手法を提案した⁶⁾。この手法は、障害物が1つの大きな塊である場合には極めて有効と考えられるものの、障害物が複数個や不特定多数個の場合には一つ一つの領域を特定するのは難しく、提案手法で想定するような街路樹や通行人などの複数の障害物をこの手法で除去することは難しいと考えられる。また、Thandaらは、ガラスなどに反射した物体をEPI上では反射層と透過層は色の線形和となって表れることを利用し、画像上から除去し、ガラスの奥の物体をはっきり映し、反射を除去した⁷⁾。しかし、実際に存在する障害物の除去手法に関しては述べてられていない。羽下らは、時空間画像解析を行い、障害物領域をエッジなどから抽出し、背景の傾きを分散で求め、その傾きで障害物領域の端から端へ滑らかに補間することで障害物除去した⁸⁾。この手法では障害物の判定をエッジ検出により行っているが、障害物のエッジが必ずしも検出できるとは限らず、また、等速直線運動を仮定した手法のため、本論文の目的としている屋外環境での等速直線運動でない場合や障害物の形状が複雑な場合に利用することが難しい。

3. 提案手法

3.1 概要

まず、車載したカメラで道路を走行しながら撮影し画像列を取得する。次に、撮影した画像を速度補正を行いEPIを作成する。作成したEPIに対して時空間画像処理により障害物を除去する。最後に、障害物を除去したEPIから、もとの画像を生成することで障害物を除去した画像を作成する。

提案手法では、EPIに対して時空間画像処理を適用し障害物を除去する際に、ヒストグラムによる補間手法を取る。このため、障害物部分の領域が背景に対して大きい場合、障害物除去に失敗することが起こる。そこで、EPI処理の前に、前処理として障害物部分のおおまかな検出を行っておく。以下でそれぞれの処理について詳細に説明する。

3.2 EPIの作成

EPIは『等速直線運動をする移動体に載せたカメラを用いて進行方向に対して垂直横向きにして撮影を行う』ことを仮定しているが、車載カメラでの撮影において、これを実現することは難しい。しかし、道路などは一般に直線状に作られており、今回提案する手法で必要な数メートル距離であれば、『直線運動』が成り立つと考えて処理することが可能である。一方で、車両で速度を一定に保つことは信号や交通状況の影響などもあり困難である。また、カメラを常に正確に進行方向に鉛直横向きに設置することも容易ではない。そこで、本論文ではこのような状況において、以下の処理により正しくEPIを作成する。

3.2.1 射影変換による画像の平行化

『進行方向に対して垂直横向き』という条件を満たすために、図1のように撮影した画像に対して射影変換にすることより平行化を行う。これによりカメラが垂直横向きに設置されていなくても、EPIを作成することができる。

3.2.2 時空間ボリュームの構築

『等速運動』という条件についてはブロックマッチングにより、移動速度を推定し、これにより時空間ボリュームの構築を行う。図2のように、ブロック初期位置 $(x, y)_{f=0} = (0, 0)$ とし、撮影した各画像間にブロックマッチングを行い、ずれ (dx, dy) を求め、 α 番目の撮影位置を $(x, y)_{f=\alpha} = (x, y)_{f=\alpha-1} + (dx, dy)$ とし、ピクセル単位で記憶する。

図3の(a)は単に画像を積層したEPI、(b)はブロック位置を考慮し構築した時空間ボリュームからのEPIである。傾きが鈍角なほど移動速度は遅く密なサンプリングを行っていることを表している。また、(b)は密な速度に合わせて補正した結果、速い移動速度の際

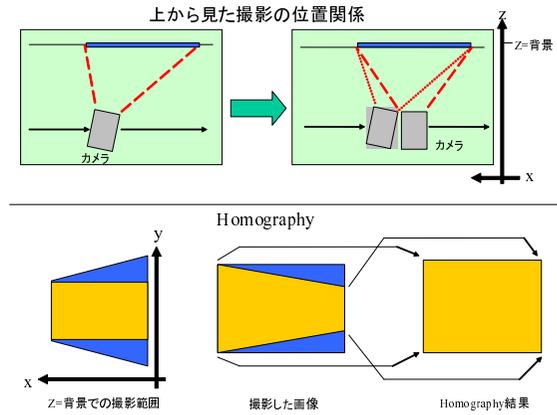


図 1 Homography 変換による平行化

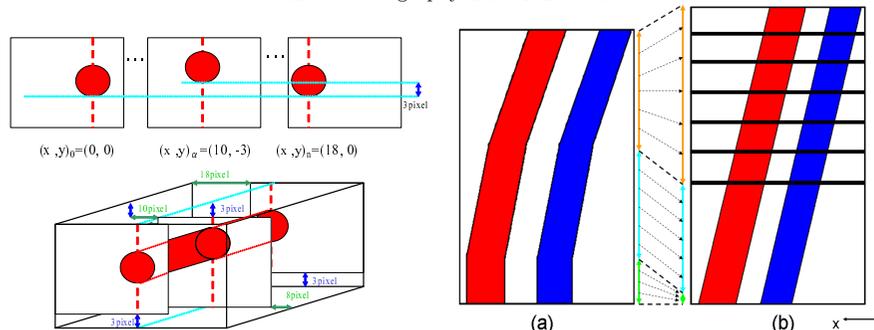


図 2 時空間ボリュームの構築

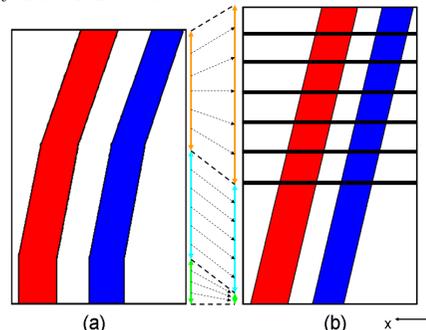


図 3 時空間ボリューム構築前後の EPI
(データの無い区間は黒で表現)

にサンプリング不足が起きていることが示されている。

3.3 障害物領域のおおまかなマスク

概要で述べた通り、背景の補間による障害物除去の前処理として、おおまかに障害物領域を検出しておく必要がある。これは、提案手法における障害物除去がヒストグラムに基づく手法のため、背景部分の出現頻度が障害物を上回り撮影されている必要があるためである。

ここで、提案手法における障害物領域の検出は正確である必要はなく、ある程度おおまかに検出できていれば良い。これは過去の障害物除去論文との大きな違いであり、これにより本手法では、街路樹などのように正確な障害物領域の検出が不可能な場合であっても、障害

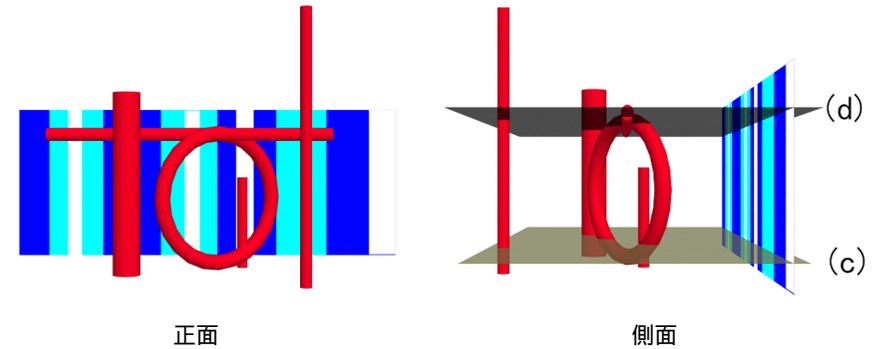


図 4 複雑なシーンの例

物除去が可能となる。図 4 のような障害物が多く存在するシーンにおける、高さ (c) の EPI を図 5 に示す。かなりの部分が障害物に覆われていることが分かる。

障害物領域の検出方法としてはオプティカルフローを用いる。具体的には、画像全体のオプティカルフローを求め、3.2.2 節で求めた撮影位置から得られる各画像間の移動距離とを比較し、それより大きいフローの部分に障害物と判定しマスクする。これは障害物の方が背景よりカメラとの距離が近いこと、画像上での移動速度が速いことを利用している。図 5 に対してマスクした結果の EPI が図 6 である。実際には、ノイズなどの影響により、この図のように障害物領域を完全にマスクすることは出来ないが、このような検出ミスは障害物除去に大きな影響を与えることはない、これは提案手法の利点の一つである。



図 5 図 4 高さ (a) の EPI



図 6 障害物をマスクした EPI

3.4 時空間フィルタによる障害物除去

時間フィルタと空間フィルタを組み合わせた時空間フィルタを使用することで障害物の除去を行う。

3.4.1 時間フィルタ

3.2.2 節で作成した図 3(b) の EPI から Canny オペレータでエッジを検出し、さらにハフ変換によりラインの傾きを求める。図 3(b) の EPI は求められた傾きのライン毎に背景のある一点を見続けていることになる。従って障害物がなければ、それぞれのライン中の画素はすべて同じとなるため、そのスキャンライン毎に背景色を検出し、その色で列全てを塗りつぶす補間を行えば障害物を除去することができる。

背景色の検出方法としては、背景の方が障害物よりも長く撮影されていれば、スキャンライン毎に色のヒストグラムを作成し最頻値を求めることで実現できる。このような前提は、3.3 節の処理により、障害物領域がおおまかに検出されマスクされていれば成立する。例として、図 6 にこの空間フィルタを適用したものが図 7 である。



図 7 時間フィルタ適用した EPI

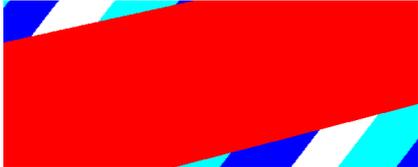


図 8 図 4 高さ (d) の EPI

3.4.2 時空間フィルタ

3.4.1 節のヒストグラムによる手法では、背景が少しでも写っており、障害物検出により、背景色が最頻値となっていれば正しく補間することができるものの、背景が一度も映っていないようなスキャンラインの場合、正しく補間することができない。このような状況は、横幅のある障害物が常に画像上にある場合やカメラと平行に移動する物体がある場合などに起こり得る。例えば、図 4 の高さ (d) における EPI がその例である。高さ (d) で生成される EPI を図 8 に示す。この例の場合には、背景が一度も見えていないため前述の手法では失敗してしまう。

そこで、各ラインごとのスキャンではなく、その前後のライン、あるいは上下の EPI 平面上のラインも考慮してヒストグラム最頻値を計算することを考える。具体的には、時空間内である程度の大きさのフィルタを設定し、そのフィルタで背景部分をスキャンしてヒストグラム最頻値を計算することで補間を行う。こうすることで、時間方向だけでは無く、空間的な補間も行われるため、図 8 のようなケースでも補間による障害物除去が可能となる。本論文では、これを時空間フィルタと呼ぶ。ここで、空間方向なフィルタサイズを大きくす

ると、補間の度合いが強まる反面、ぼけにより画質が劣化してしまう。そこで、等方的ではないフィルタを利用することで、このような影響を限定することができる。例えば、除去の対象物が電線のような場合、電線は細くほぼ水平に位置しているため、縦方向に長いフィルタを利用すれば画質の劣化を抑えながら、電線の除去が可能となる。

3.5 背景の奥行きレベルが複数の場合

これまでは背景が一つの平面と仮定して処理を行ってきた。しかし、実際には例えば市街などで隣接するビル奥ゆきが同じであるとは限らず、背景に複数の奥行きを持った平面が存在することがある。本論文ではこのような複数デプスの場合にも影響されずに背景を補間する方法を提案する。

複数の平面を仮定したシーン図 9 の高さ (e) における EPI を図 10 に示す。Canny オペレータとハフ変換によるエッジ検出で求めた傾き毎にラインの両端を求め、そのラインで挟まれた領域がデプス毎の背景領域であり、色分けして示すと、図 11 のようになる。それぞれの領域とその傾き毎に補間処理をすることで背景が複数デプスの場合でも障害物を除去することができる。本論文では、デプスごとの領域検出に関しては、実画像での自動処理が困難であったため、Canny オペレータとハフ変換によるエッジ検出後、手動で行った。領域検出の自動化は今後の課題である。

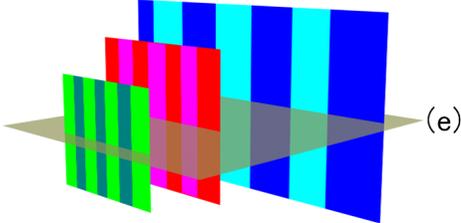


図 9 奥行きレベルが複数の例



図 10 複数デプスの EPI



図 11 複数デプスの背景領域 (デプスごとに処理を行う)

4. 実験

4.1 評価実験：室内

本手法の評価のため、室内でモーターを使用しカメラを等速直線運動で動かし撮影した(図12)。撮影したビデオ映像を用いて、障害物の除去および真値との比較実験を行った。撮影した動画からキャプチャーした画像が図13であり、この画像列から作成したEPIが図14である。これに対して、デプス毎に障害物のマスクと空間フィルタを適用すると図15のようになる。背景の奥ゆきが複数の場合でも障害物が除去できたことが確認できる。障害物を除去して撮影した画像を用いて計算するとPSNRは20.3dbであった。



図12 複数デプス撮影風景



図13 複数デプスカプチャー画像



図14 図13の300pixel目の複数デプスEPI



図15 図14に空間フィルタを適用した複数デプスEPI



図16 複数デプス除去結果画像

4.2 屋外実験1：レールによる実験

屋外でレールとドリーを使用し、レール上をカメラを手押しで進め、撮影した。撮影した動画からキャプチャーした画像が図17である。この場合、カメラとの距離が近い垣根が障害物であり、奥の壁の下半分がほとんど撮影されていない。撮影した映像にブロックマッチングを行って時空間ボリュームを構築し300pixel目の高さにおけるEPIが図18である。また、障害物は幅だけでなく奥行きもあり撮影位置によって常にデプスが変わりEPI上での変化が激しい。図18にそのまま補間処理を行うと図19になり、障害物が背景より支配的なため、ライン毎のヒストグラムピークで障害物の色が選択されてしまい背景色の選択に失敗してしまう。

そこで、障害物をオプティカルフローによりマスクした画像が図20である。図20からEPIを作成すると図21となる。障害物にマスクをしてしまったため色情報がほとんどないように見えるが、理論的には、列中に背景のみが1pixelでも残っており他に画素がなければ採用されるため、これに時空間フィルタを適用すると図22になる。この処理をすべての高さにおけるEPIについて行い、元の画像に戻すと図23になり、垣根が除去できたことが確認できる。

さらにフェンスのあるシーンにおいても障害物除去を行った。入力と結果を図24に示す。フェンスのように細い障害物の除去は一般に困難であるが、正しく除去することができていることが確認できる。

4.3 屋外実験2：車載映像による実験

図25のような全方位カメラLadybug2¹²⁾を計測車両に搭載し、走行しながら撮影を行った。全方位カメラだと、あらゆる方向の画像が一度に撮影できるため、同じ物体(建物)が長く撮影されることになり、時空間画像解析にとって都合が良い。

撮影した全方位画像を図26に示す。図26を透視投影変換したものが図27である。図27



図 17 入力キャプチャー画像



図 18 図 17 の 300pixel 目の EPI

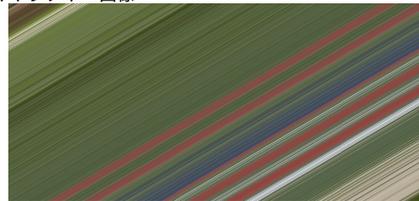


図 19 障害物除去に失敗した EPI



図 20 障害物をマスクした画像

の 300pixel 目の高さ、および、ちょうど電線部分の高さ (f) における EPI が、それぞれ図 28 および図 30 である。図 30 を見ると、電線の影響により、EPI 全体に不規則なパターンが観測される。これに空間フィルタを適用した結果が図 29 および図 31 である。図 31 では、白い柱の部分に黒い色が混じるなど、うまく電線が除去できていないことが分かる。この

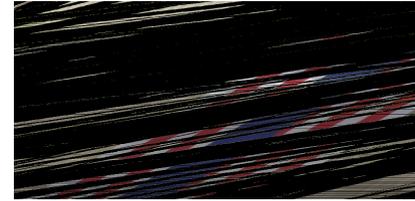


図 21 障害物をマスクした EPI



図 22 空間フィルタを適用した EPI



図 23 障害物除去結果画像

障害物除去後の EPI を画像に戻すと、図 32 となり、電柱は削除できているものの、電線部分が残っていることが分かる。

これに対して、図 30 に対して、空間フィルタのサイズを縦一列 13pixel とした時空間フィルタを適用した EPI が図 33 である。先ほどと異なり、黒いノイズなどが無くなっていることが分かる。この結果復元された画像が図 34 であり、時空間フィルタの使用により、電柱と電線のそれぞれが除去できたことが確認できる。

5. ま と め

本論文では時空間解析の一つである EPI 解析を利用して障害物の除去に関して述べた。まず最初に、EPI を作成するにあたり、等速でない状況やカメラの上下振動によるのずれに対応し、正しい EPI を作成する手法、および、オプティカルフローにより障害物をおおまかに除去した EPI を作成する手法を紹介した。次に、こうして得られた EPI に対して、時空間フィルタによりヒストグラム最頻値による補間を行うことで、効率よく障害物を除去す



図 24 フェンス除去



図 25 Ladybug2 と計測車

る手法を提案した．最後に，屋内・屋外で撮影したデータを用いて実験を行ったところ，垣根などの複雑な形状や，移動する物体について正しく障害物除去できることを確認できた．

参 考 文 献

- 1) 小幡 恭久, 剣持 雪子, 小谷 一孔, "画像のフラクタル性を活用した局所的な画像推定法による画像復元手法" PRMU2002-126 pp. 37-42
- 2) 榎本 暁人, 斎藤 英雄, "複数のハンディカメラを利用した Diminished Reality", MIRU2007
- 3) Yin Li, Jian Sun, Chi-Keung Tang and Heung-Yeung Shum, "Lazy Snapping", SIGGRAPH 2004, Vol. 23, pp. 303-308.



図 26 入力全方位画像

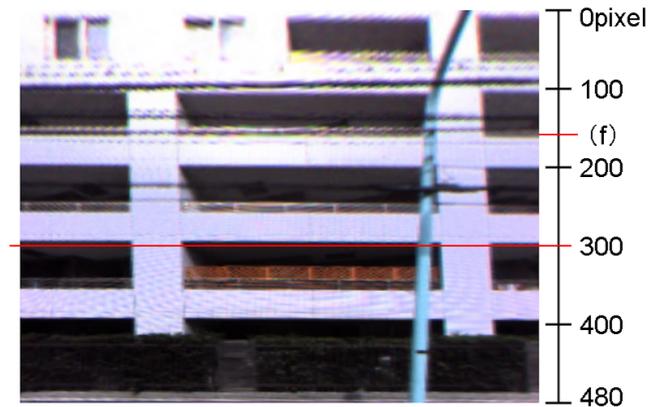


図 27 透視投影変換画像

- 4) C. Rasmussen, T. Korah, "Spatiotemporal inpainting for recovering texture maps of partially occluded building facades" In Proceedings of the IEEE ICIP.
- 5) Yasuyuki Matsushita, Ko Nishino, Katsushi Ikeuchi, Masao Sakauchi, "Illumination Normalization with Time-Dependent Intrinsic Images for Video Surveillance", IEEE Trans. Pattern Anal. Mach. Intell., 26(10): 1336-1347 (2004).
- 6) Takaaki Shiratori, Yasuyuki Matsushita, Sing Bing Kang, and Xiaoou Tang, "Video completion by motion field transfer", In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), volume 1, pages 411-418, June 2006.
- 7) Thanda Oo, 川崎洋, 大澤裕, 池内克史, "Separation of Reflection and Transparency Based on Spatiotemporal Analysis for Outdoor Scene", IPSJ Vol. 2(2006) pp.428-440.



図 28 図 27 の 300pixel 目における EPI



図 29 図 28 に空間フィルタを適用した EPI

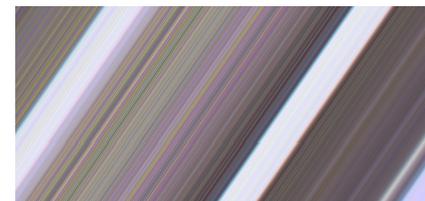


図 33 時空間フィルタを適用した電線の高さ (f) の EPI

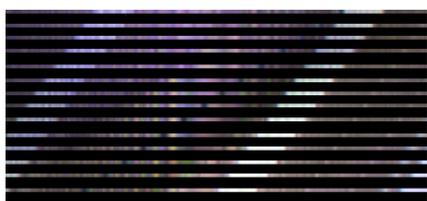


図 30 電線の高さ (f) の EPI



図 31 時間フィルタを適用した電線の高さ (f) の EPI



図 32 空間フィルタを適用した除去結果画像



図 34 時空間フィルタを適用した除去結果画像

- 8) 羽下哲司, 鷺見和彦, 橋本学, 関明伸, 黒田伸一, "時空間画像処理による注目距離対象の強調に基づいた遠隔監視システム", 電子情報通信学会技術報告 (PRMU97-126), pp.41-46 (1997).
- 9) 三上 武志, タンダー ウー, 小野 晋太郎, 川崎 洋, 大沢 裕, 池内 克史" EPI 解析を利

- 用した歪みのない複数ビデオカメラ画像の総合" 電子情報通信学会論文誌 D-II, No.6, pp.1336-1347, 2006
- 10) 王金戈, 小野晋太郎, 池内克史, "一次元メディアンフィルタを用いた市街地画像からの電線除去手法の提案", 第 7 回 ITS シンポジウム, Dec. 2008. pp.201-203.
- 11) 永橋知行, 藤吉弘巨, 金出武雄, "Spatio-temporal volume を用いた繰り返し処理のグラフカットによるビデオセグメンテーション", MIRU2008.
- 12) P.G.R.Inc.:Ladybug : "http://ptgrey.com/."