

直交視点画像を用いた3次元流体モデリング

Modeling Fluids from Ortho-view Images by Appearance Transfer

岡部 誠¹ 土橋 宜典² 山口 尊嗣³ 安生 健一⁴ 尾内 理紀夫⁵

Makoto Okabe¹ Yoshinori Dobashi² Takatsugu Yamaguchi³ Ken Anjyo⁴ Rikio Onai⁵

1,3,5 電気通信大学

1,3,5 The University of Electro-Communications

2 北海道大学

2 Hokkaido University

4 オール・エル・エム・デジタル / 科学技術振興機構 CREST

4 OLM Digital, Inc. / JST CREST

E-mail: 1 m.o@acm.org 2 doba@ime.ist.hokudai.ac.jp 3 yamaguchi@onailab.com 4 anjyo@olm.co.jp 5 onai@cs.uec.ac.jp

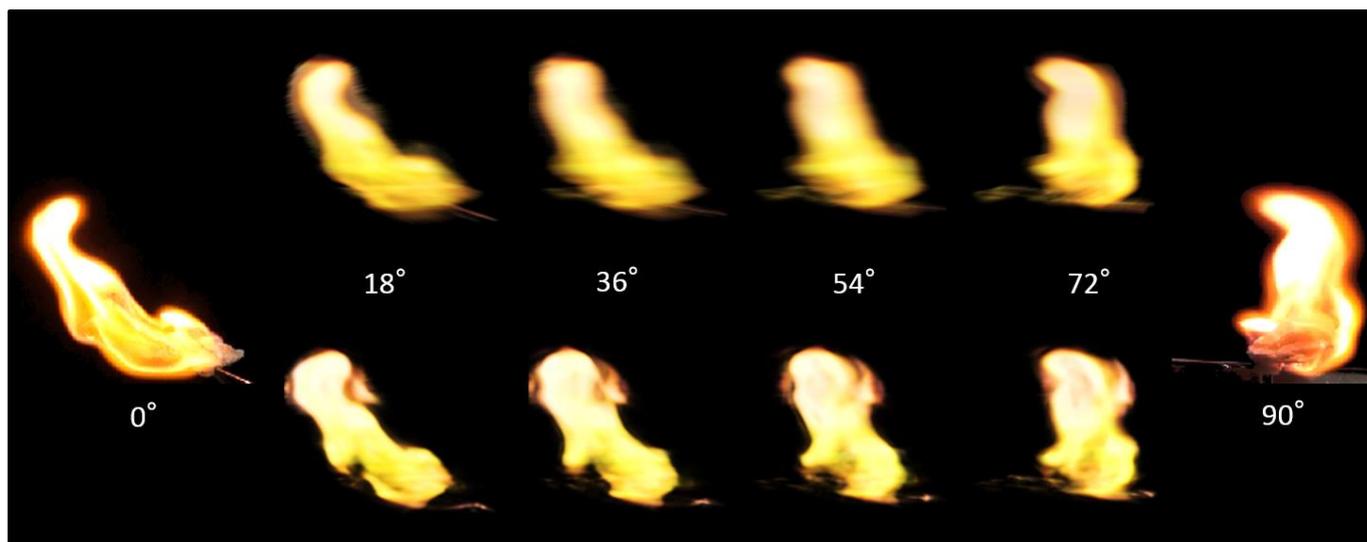


図 1: 左端と右端は実際の炎を正面と真横から撮影した画像である。上段はこれらの画像から既存手法で作ったボリューム。回転して別の視点から見ると不自然にボケて見える。下段は提案手法で作ったボリューム。別の視点から見た時の見た目が保たれるようにボリュームを作るので、どこから見ても入力画像の見た目に近いボリュームが出来る。

1. はじめに

炎、煙、爆発など、流体の視覚効果は映像製作に欠かせない。実際、コンピュータ・グラフィックスの研究分野では、流体は常に最も重要な課題の1つとして捉えられてきた。映像製作の現場では、流体のシーンは大きく2つの方法で作られている。1つは流体シミュレーションによる方法[1]、もう1つは流体を撮影した動画を合成する方法である。本論文は後者の問題を扱う。

多くの映像製作会社は独自の動画データベースを持っている。ポストプロダクションを行うアーティストは、データベースから適切な流体動画を選び、動画編集ツールでシーンに重ね合わせる。しかし、動画は2次元なので、カメラが動くシーンや立体映像の制作など、3次元的な効果が必要な場合には対応できない。また多くの流体は半透明なので、自然に合成するために複雑なマットが必要な場合は大変時間の掛かる作業となる。

2次元流体動画の欠点を補うため、ユーザが少ない労力で3次元流体を作ることのできるイメージ・ベース

な手法を提案する。提案手法の入力は2枚の直交視点画像のみである(図1左端と右端)。そのため、流体撮影時のスタジオ・セットアップが簡単である。既存のイメージ・ベースの手法では、多くのカメラを流体の周囲に配置する必要があった[2,3,5]が、提案手法では2つのカメラがだいたい直角になるように配置できれば良く、カメラ・キャリブレーションも不要である。動画が与えられた後は、提案手法が自動的に3次元ボリュームを作る。

入力画像が2枚なので、既存のイメージ・ベースの手法を使うには、3次元ボリュームの構築に必要な情報が不足している。実際、図1上段に示すのは、Ihrkeらの手法を適用して作ったボリュームだが、回転して別の視点から見ると不自然にボケて見える。我々が何故これをボケていると感じるか、と考えると、我々は2つの入力画像(図1左端と右端)の見た目を既に知っていて、それらと比較するのでボケていると感じるのではないかと、というのが我々の仮説である。提案手法

はこの仮説に基づいている．2枚の直交視点画像のような見た目になるように，別の視点から見た時の画像を修正する．修正された画像を用いて3次元ボリュームを更新する．この修正と更新を繰り返すと，ボリュームはより良い見た目を持つものに収束していく．図1下段にその結果を示す．

今回の提案は，画像から3次元ボリュームを正しく再構築するための手法ではなく，あくまで3次元ボリューム・モデリングの手法である．2枚の画像のみから，そこに写る3次元ボリュームを正確に再構築するのは非常に難しい問題である．一方で，我々は2枚の画像のみからでも，映像製作に使えるような質の3次元ボリュームが作れることを示す．従って，入力画像は再現すべき目標ではなく，あくまでモデリングのための参照画像という位置付けである．

2. 既存研究

既存の computed tomography 技術(CT)は，多視点で撮影された X 線画像が与えられれば，そこに写る3次元半透明物体を再構築できるという，Radon の投影理論に基づいている[7]．だが，そのような CT スキャナを

炎や煙の撮影に用いるのは難しいため，画像ベースの tomography 技術が提案されてきた．

Ihrke らは多視点で撮影した炎や煙の画像から最小二乗法による最適化を用い，3次元ボリュームを作る手法を提案した[5]．Hasinoff らは炎の画像生成モデルと flame sheet という炎の表現法を提案し，更に flame sheet の線形和で複雑な炎を表現する手法を提案した[3]．Gregson らは統計的最適化に基づく CT を提案した[2]．統計的手法を用いることで見た目に関する正規化項など，高レベルで複雑な制約をエネルギー関数に含めることが可能となった．これらの手法は高品質な流体の3次元ボリュームを作ることができるが，密に撮影された多視点画像が必要であり，非専門家が簡単に扱える手法ではない．

3. 提案手法

2枚の直交視点画像(図2-a)が与えられると，Ihrke らの手法(LSM=最小二乗法)を用いて初期ボリュームを作る(図2-b)．このボリュームは図1上段のように別の視点から見ると不自然にボケていたり，図2-bのように斜め上から見下ろすと不自然なグリッドが見える．

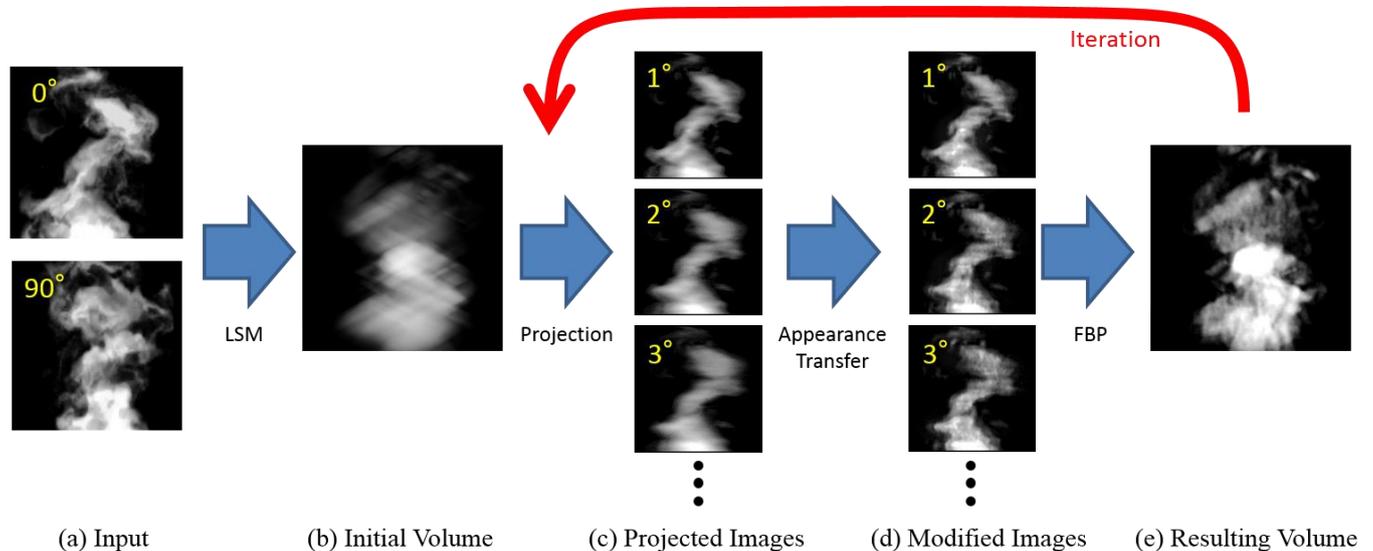


図2: 提案手法の処理の流れ．

提案手法は次の3ステップを繰り返す．1) 現在のボリュームを多視点方向へ投影してレンダリングする(図2-c)．2) テクスチャ合成手法を用い，各投影画像を入力画像と似たような見た目になるよう修正する(図2-d)．3) 修正した画像にフィルタ補正逆投影法を適用し，3次元ボリュームを構築する(図2-e)．3ステップ終了後，再びステップ1へ戻り(図2の赤い矢印)，構築されたボリュームは多視点方向へ投影される．

以下，入力画像はグレースケールとし，流体の密度を表す3次元ボリューム(同じくグレースケール)を構築する手法を述べる．初期ボリュームは3.1章，3ステップは3.2章から3.4章で述べる．以上を基本アルゴリズムとし，それを3.5章で多重解像度に拡張する．

3.1. 初期ボリューム

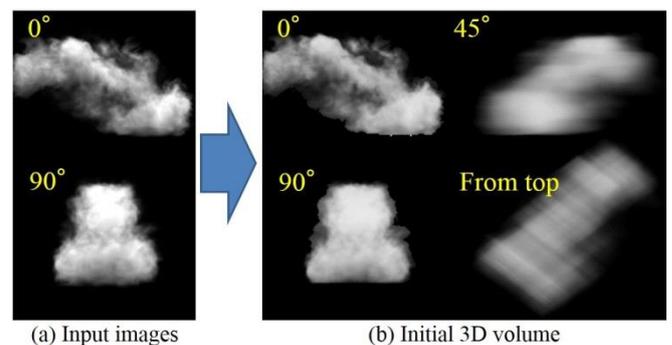


図3: 煙のCG画像から求めた初期ボリューム．

Ihrke らの最小二乗法に基づく方法を用い、2枚の直交視点画像から初期ボリュームを作る。求めたいボリュームを Ray Casting 法によって正面及び真横方向に投影したものが入力画像だと仮定すると、 $O = BV$ が成り立つ。 O は入力画像のピクセル集合、 B は Ray Casting 法の処理に相当する行列、 V は求めたいボリュームである。 O と B が与えられているので、 V は擬似逆行列を用いて、 $V = (B^T B)^{-1} B^T O$ 、と書ける。共役勾配法を用いて計算すれば V が求まる。

図 3-a は煙のシミュレーションデータをレンダリングして作った直交視点画像、図 3-b はそれらに Ihrke らの手法を適用して作った初期ボリュームである。正面(0度)及び真横(90度)から見ると入力画像に似ているが、斜め45度から見ると不自然にボケている。また、斜め上から見下ろすとグリッド模様が見える。

3.2. 多視点で投影

ステップ1では、現在の3次元ボリュームを正射影で多視点に投影した画像を作る。この投影には Ray Casting 法を用いる。投影方向は $[0, \pi)$ の範囲で、角度の間隔は1度、従って180枚の投影画像を作ることとなる。図4は図3の初期ボリュームを15度、45度、75度で投影した画像である。

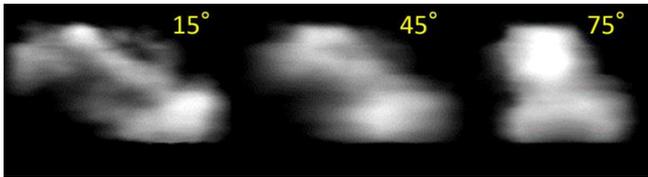


図4: 初期ボリューム(図3)の投影画像。

3.3. 見た目の復元

ステップ2では、入力画像に似た見た目を持つように投影画像を修正する。画像のピラミッドとヒストグラム・マッチングを用いた Heeger らのテクスチャ合成法を使う[4]。本手法では steerable pyramid を用い、画像を4スケール、4方向の帯域に分解する(図5)。投影画像を修正するには、入力画像と投影画像を共に steerable pyramid で分解し、対応する帯域間で、投影画像の帯域が入力画像の帯域と同じヒストグラムを持つようにヒストグラム・マッチングを行う。修正された steerable pyramid から画像を復元すると、入力画像に似た見た目の投影画像が得られる、というのが Heeger らの手法である。

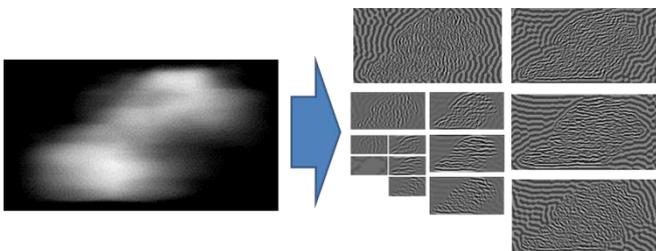


図5: 4スケール、4方向成分を持つ steerable pyramid.

入力画像は正面と真横の2つなので、そのどちらかの帯域のヒストグラムを使うのではなく、両者のヒストグラムを線形補間してマッチングに使う。図6-aとcは正面(0度)と真横(90度)の入力画像の帯域である。図6-dとfはそれらのヒストグラムである。それぞれ b^0 、 b^{90} とする。今、 x 度での投影画像を修正したいとすると、目的のヒストグラム b^x の各ビンを次の式で線形補間して求める(図6-e):

$$b^x(i) = \frac{90-x}{90} b^0(i) + \frac{x}{90} b^{90}(i).$$

ここで i はヒストグラムのビンのインデックスである。 x 度での帯域のヒストグラムが、 b^x に等しくなるように帯域を修正する(図6-b)。

図7は図4の投影画像それぞれにつき、見た目を復元した画像である。入力画像に近く、はっきりした見た目となっている。

シミュレーションで作成した炎のボリュームから180枚の投影画像を作り、それらの帯域のヒストグラムが $[0, \pi)$ の範囲でどのように変化するか観察した。ヒストグラムの大きな形状はスムーズに変化した。一方、その変化は必ずしも線形ではなかった。今回、線形補間を採用したのは、実装が簡単でスムーズな補間が実現できるからだが、線形補間よりも優れた補間方法が存在する可能性はあり、調査は将来の研究課題である。

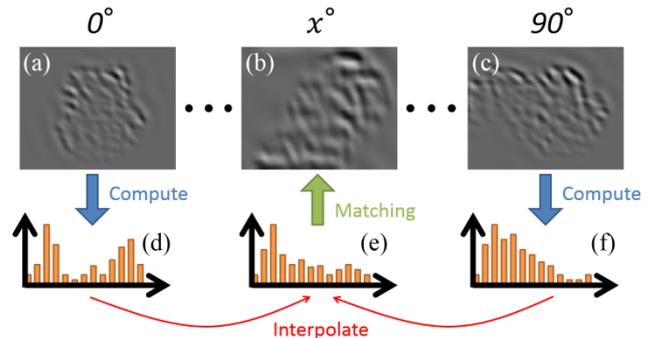


図6: ヒストグラムの線形補間とマッチング。

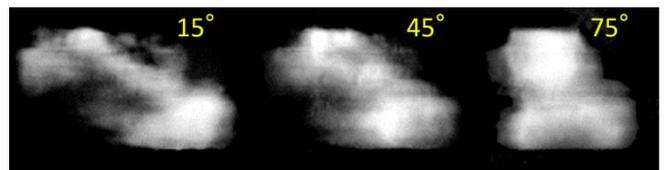


図7: 投影画像(図4)の見た目を復元した結果。

3.4. フィルタ補正逆投影法の適用

ステップ3では、フィルタ補正逆投影法を適用して修正した投影画像(図7)から3次元ボリュームを再構築する。ラドンの定理から導かれるフィルタ補正逆投影法を使えば、投影画像にフィルタを掛けた画像を加算処理することで、他の手法[2,3,5]と比べて高速にボクセル値を決定することができる[7]:

$$V(x, y, z) = \sum_{\theta=0}^{179} I^{\theta}(x \cos \theta + y \sin \theta, z) * FT^{-1}(|f|),$$

I^{θ} は角度 θ での投影画像、 $*$ は畳み込み演算、 FT^{-1} は逆フーリエ変換、 f は周波数を表している。

3.5. 多重解像度への拡張

2次元及び3次元テクスチャ合成の既存手法では、低解像度なものから高解像度なものへ順次合成処理を繰り返すような多重解像度を用いたアルゴリズムが、高画質なテクスチャ画像を合成できることが知られている[6]。我々も3.1章から3.4章で述べた基本アルゴリズムを多重解像度に拡張する。入力画像の低解像度版を作り、低解像度なボリュームを基本アルゴリズムで合成する。次に入力画像及びボリュームの解像度を現在の2倍にし、再び基本アルゴリズムを適用する。この処理を最大解像度まで繰り返す。この過程は付録のビデオで示した。

図8に多重解像度アプローチの利点を示す。図8-aは入力画像である。図8-bと図8-cは多重解像度アプローチを用いた場合と用いない場合の結果を示す。単一解像度で作ったボリュームは尖がっていてノイズが見える。3段階の多重解像度アプローチを用いた方は滑らかで丸い形で作られている。

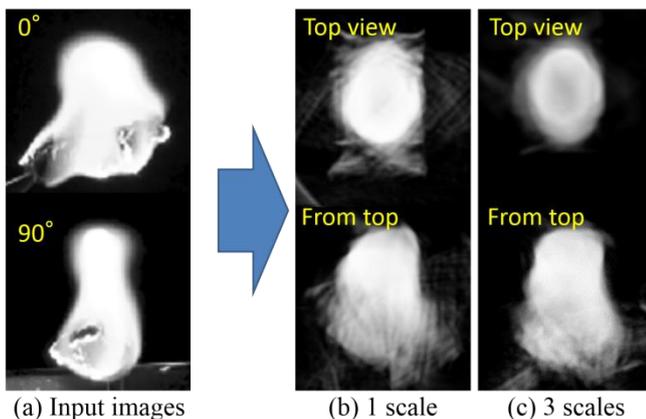


図8: 多重解像度アプローチの効果

4. 結果と考察

提案手法をCGで合成した入力画像と実際に撮影した入力画像の両方に適用した。付録のビデオに3次元ボリュームのアニメーションを示す。

図1の炎は、紙を燃やした際の炎をソニー・ハンディカム2台で撮影している。カメラ・キャリブレーションは行わず、代わりに動画処理ソフトウェアで2つの動画のタイミング、位置、サイズを手動で合わせた。動画を与えると、提案手法は全自動で炎の3次元ボリュームを合成した。既存手法では視点を変えた時に見た目がボケたり、付録のビデオに示すように斜め上から見るとグリッドが見えるが、提案手法はどこから見

ても丸い炎の形を作ることに成功している。一方、図1下段の画像を入力画像と比較すると、まだ鮮明さが足りず、また入力画像にない模様も見えるので、入力画像の炎が正確に再構築できた訳ではない。しかし、合成された3次元ボリューム自体の質を評価するならば、ある程度の実用にも耐えられる3次元の炎ができていると考えられる。

図9は図3左の入力画像から得られた結果である。図3右と比較すると、入力画像のモコモコした煙の雰囲気を得られていると共に、グリッドも消えている。

図10の炎も、シミュレーションで得たボリュームデータを正面及び真横からレンダリングして2枚の入力画像を作り、提案手法でボリュームを作ったものである。図10-eと図10-f(及び付録のビデオ)で45度の角度からの見た目を正解のボリュームデータの見た目と比較している。提案手法の結果はやや鮮明さに欠け、正解データを正確に復元できている訳ではないが、炎のテクスチャや全体的な雰囲気が再現できている。

2枚の入力画像の両方に同じ画像を使えば、単一画像のみからでもある程度の3次元ボリュームを作ることができる。図11は単一の煙の動画から作成したボリュームのアニメーションである。付録のビデオに示すように、単一画像から作ったボリュームはある視点から見ると左右対称になる。図11の各フレームの左側は入力画像、右側は光の1回散乱を考慮したボリューム・レンダリングの結果である(視点は入力画像と異なる)。

図12は爆発の単一画像から作成したボリュームのアニメーションである。爆発の丸い形やテクスチャが3次元的に再現できている。付録のビデオでは入力画像の片方に骸骨の画像を重ね合わせて、骸骨の顔を持つ爆発ボリュームを作った例も載せている。

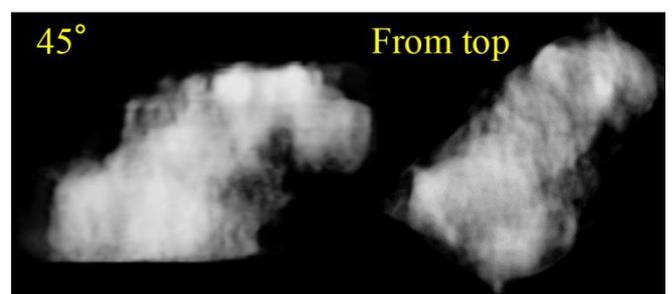


図9: 図3の初期ボリュームから作った結果。

構造を保持できないという限界 提案手法を用いて2枚の樹木の直交視点画像から3次元の樹木を作る実験を行った。炎や煙と同じ画像生成モデルが樹木に対しても適用できる事が既存研究で提案されており、それを用いたものである[8]。しかし、提案手法は樹木の枝構造を作ることに失敗した。図13に結果のボリュームを示すが、付録の動画にも示すように、樹木の正面(もしくは側面)から少しでも視点をずらすと、図13の赤い楕円の中の幹が消滅してしまう。これは、今回

用いたテクスチャ合成手法[4]が樹木の枝など、長いストロークのような構造を表現する事に向いていないために起こる現象であり、幹はボリューム内部に散らばっているためである。この問題を解決するために、将来、別のテクスチャ合成手法を調査し、適用の可能性を探りたい。例えば、Portillaらの手法は統計的なテクスチャ合成手法でありつつ、画像の構造をより良く表

現できる手法を提案している[9]。

解像度と計算時間 図3及び図9の煙の例は、入力画像の解像度は 188×150 、従って生成されるボリュームのサイズは $188 \times 188 \times 150$ となる。このような比較的小さな解像度のボリュームを作る場合でも提案手法は3分50秒の時間を要する。ボトルネックは3.3章の見た目の復元である。高速化の可能性を探りたい。

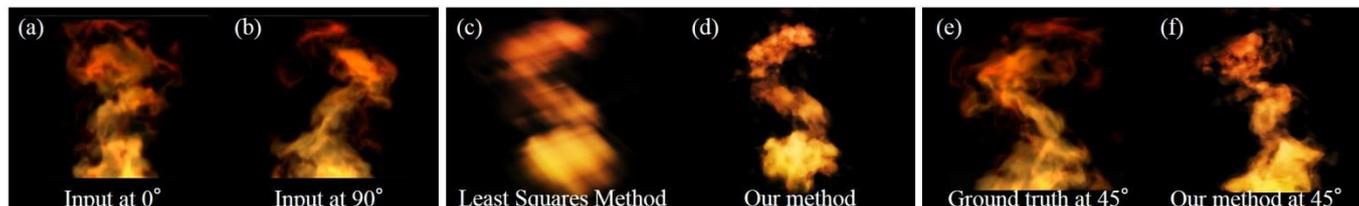


図 10: 炎の CG 画像を用いたボリューム・モデリングの結果。

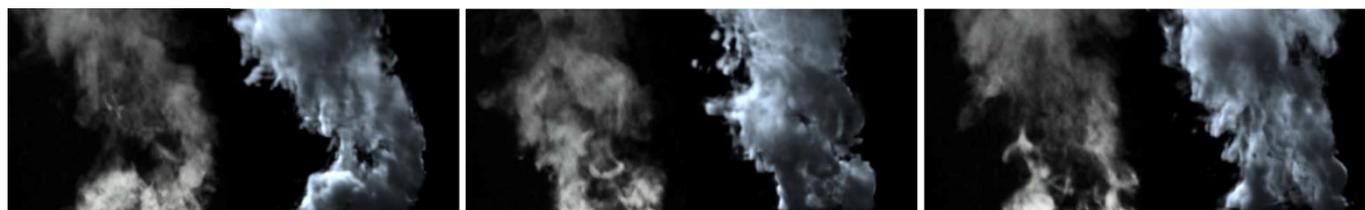


図 11: 単一のカメラで撮影された煙の動画を用いたボリューム・モデリングの結果。

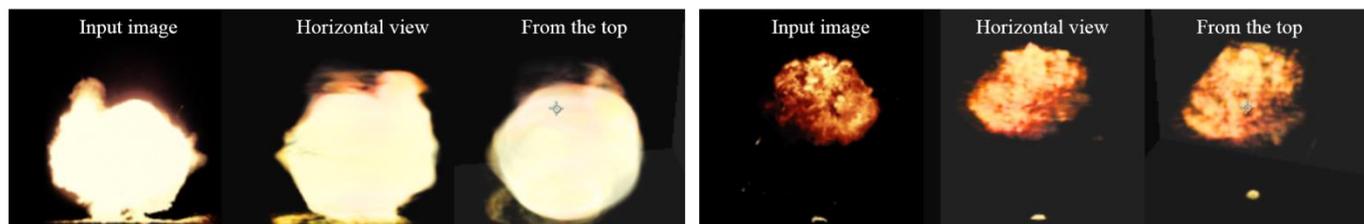


図 12: 単一のカメラで撮影された爆発動画を用いたボリューム・モデリングの結果。

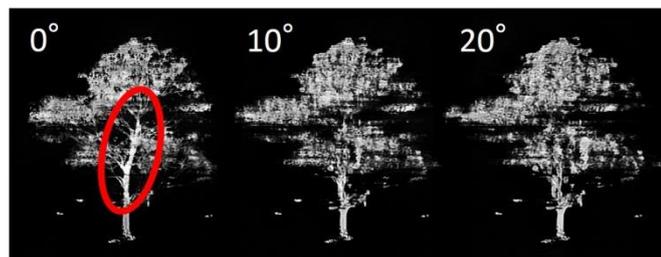


図 13: 提案手法で作った樹木の 3 次元ボリューム。赤い楕円の中の幹は視点を変えるとすぐに消えてしまう。

参考文献

- [1] Bridson, R., and Müller-Fischer, M. 2007. Fluid simulation: Siggraph 2007 course notes. In ACM SIGGRAPH 2007 Courses, 1–81.
- [2] Gregson, J., Krimerman, M., Hullin, M. B., and Heidrich, W. 2012. Stochastic tomography and its applications in 3d imaging of mixing fluids. ACM Trans. Graph. 31, 4, 52:1–52:10.
- [3] Hasinoff, S. W., and Kutulakos, K. N. 2007. Photo-consistent reconstruction of semitransparent scenes by density-sheet decomposition. IEEE Trans. Pattern Anal. Mach. Intell. 29, 5, 870–885.
- [4] Heeger, D. J., and Bergen, J. R. 1995. Pyramid-based texture analysis/synthesis. In Proc. of SIGGRAPH '95, 229–238.
- [5] Ihrke, I., and Magnor, M. 2004. Image-based tomographic reconstruction of flames. In Proc. of SCA '04, 365–373.
- [6] Kopf, J., Fu, C.-W., Cohen-Or, D., Deussen, O., Lischinski, D., and Wong, T.-T. 2007. Solid texture synthesis from 2d exemplars. ACM Trans. Graph. 26, 3.
- [7] Schuster, T. 2007. A filtered back projection algorithm. In The Method of Approximate Inverse: Theory and Applications, vol. 1906 of Lecture Notes in Mathematics, 165–179.
- [8] Reche-Martinez, A., Martin, I., and Drettakis, G. 2004. Volumetric reconstruction and interactive rendering of trees from photographs. ACM Trans. Graph. 23, 3, 720–727.
- [9] Portilla, J., and Simoncelli, E. P. 2000. A parametric texture model based on joint statistics of complex wavelet coefficients. Int. J. Comput. Vision 40, 1, 49–70.
- [10] McCann, J., and Pollard, N. S. 2012. Soft stacking. Comp. Graph. Forum 31, 2pt2 (Eurographics), 469–478.