



奥行推定と画像領域分割の融合による デプスマップの精度向上技術

NTTサービスエボリューション研究所では、視聴覚に働きかけ情報を効果的に提示することにより、ナチュラルなインタラクションを実現する新たなメディア処理技術の研究開発を進めています。本稿では、奥行推定と画像領域分割の融合によるデプスマップの精度向上技術と、本技術を用いて単眼2D映像からナチュラルな3D視聴を実現するHiddenStereo映像を生成するシステムを紹介します。

おの 小野	まさと 正人	きくち 菊地	ゆみ 由実
さの 佐野	たかし 卓	ふかつ 深津	しんじ 真二

NTTサービスエボリューション研究所

奥行推定と画像領域分割の融合による デプスマップの精度向上技術の概要

デプスマップ（深度マップ）とは、画像の各画素に対してカメラからの距離を表現したものであり、さまざまな用途で活用されています。身近な例では、スマートフォンで写真撮影をした際の、距離が遠い背景をぼかすような写真加工や、自動運転の分野における車両周辺の物体検出などに活用されています。

私たちは、「エンタメ分野における3D表現」をターゲットとし、新たに3D撮影するもの

だけでなく、既存の2Dコンテンツも3Dコンテンツ化可能とするためのメディア処理技術の研究開発に取り組んでいます（図1）。本技術は、2D映像から精度の高い奥行き情報を生成する「デプスマップ生成技術」と、効果的な3D表現等のためにデプスマップを補正する「デプスマップ最適化技術」から構成されます。

技術のポイント

■デプスマップ生成技術

デプスマップを得る手法としては、ステレ

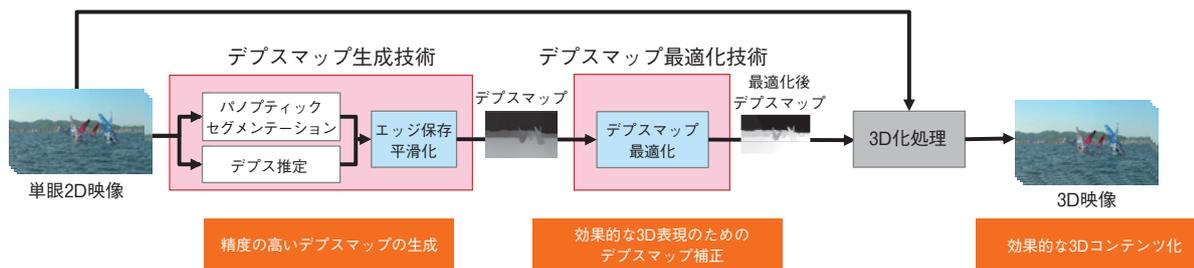


図1 奥行推定と画像領域分割の融合によるデプスマップの精度向上技術の概要

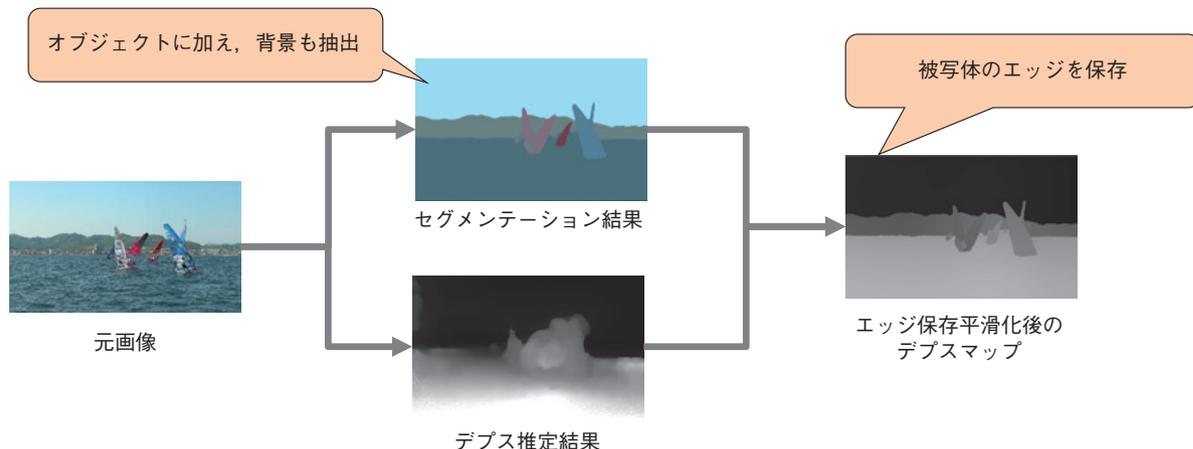


図2 デプスマップ生成技術の概要

オカメラにおけるカメラ間の視差を活用する手法、カメラ映像と外部機器（LiDAR等）の情報を組み合わせる手法、各種画像処理により生成する手法などが挙げられます。近年、既存の2Dコンテンツからデプスマップを生成する際に用いられる手法としては深層学習を活用するものがありますが、一般的に生成されるデプスマップの解像度は低く、また被写体の輪郭がクッキリとしたデプスマップを得ることはできませんでした。ここで、4K・8Kといった高解像度なアーカイブ映像に適用する場合、デプスマップも高解像度かつ品質の高いものを得る必要があります。

そこで、デプスマップの補正手法として提案されている、原画を用いたエッジ保存平滑化手法^{(1),(2)}を参考に、映像中の各被写体の輪郭をよりクリアにし、効果的な3D表現を実現する手法として、「セグメンテーション結果を用いたエッジ保存平滑化」による、デプスマップの精度向上手法（図2）を提案します。具体的には、デプス推定で得た解像度の低いデプスマップに対して、セグメンテーション結果画像をガイドとするフィルタ処理を行うことにより、デプスマップの解像度を向上

しつつ、被写体に含まれるオブジェクトのエッジを保存することを可能とします。このとき、セグメンテーション結果としては、オブジェクトに加え背景部分も各要素で分離できるよう、セマンティックセグメンテーションとインスタンスセグメンテーションを足し合わせたパノプティックセグメンテーションを利用しています。

本手法においては、デプス推定およびセグメンテーションは疎な関係となっているため、各処理のアルゴリズムは適宜変更することが可能です。例えば、現状ではデプス推定およびセグメンテーションには深層学習のアルゴリズムを用いていますが、その他のアルゴリズム、例えば自己教師あり学習のアルゴリズムなどに交換することも容易です。

■デプスマップ最適化技術

前述したデプスマップ生成技術により、精度の高いデプスマップを得ることが可能となりますが、デプスマップをそのまま利用した場合には、必ずしも利用用途に合致した内容となっていないケースがあります。例えば3D映画等では、単純に映像内の奥行きをデプスマップとして表現できる奥行の範囲に一

律で縮小すると、映像内の奥行き差にメリハリが出ず、効果的な3D表現とはなりません。

これに対し、3D映画等では、映像視聴を行う際に臨場感が向上するように、デプスマップを補正する手法が活用されています。そこで私たちも、注目すべき被写体周辺の3D感をより強調し、効果的な立体感を得られるよう、デプスマップを最適化する手法（図3）を提案します。具体的には、2つの観点での補正を行うものとなります。

(1) 3D表現に利用する奥行き範囲の限定

3D視聴時の視覚的な傾向として、注視する被写体から奥行きが大きく異なる部分（遠方の背景や、手前に映り込んだ物体など）は、細かい奥行き差が区別しにくいいため、そのような範囲に対しては割り当てる奥行き差の範囲を減らします。具体的には、0から255までの256段階で表現されるデプスマップをヒ

ストグラム解析し、被写体が存在する範囲、例えばデプスマップのデプス値の分布から算出した5パーセンタイルから95パーセンタイルの範囲を「有効なデプスの範囲」と定め、その範囲に含まれるデプス値を0から255の値に伸長し、それ以外の範囲は0（もっとも奥）または255（もっとも手前）にマッピングします。このことにより、3D視聴時に有効なデプスの範囲のみ3D表現することを可能とします。

(2) 奥行き差の強調表現

映像内の注視する被写体がある範囲の奥行き感を強調表現することにより、よりメリハリのある3D映像を生成します。具体的には、デプスマップのヒストグラム解析結果から、注視すべき被写体が存在するデプスの範囲をデプスレイヤとして導出し、もっとも注目したい被写体が存在するデプスレイヤの幅を

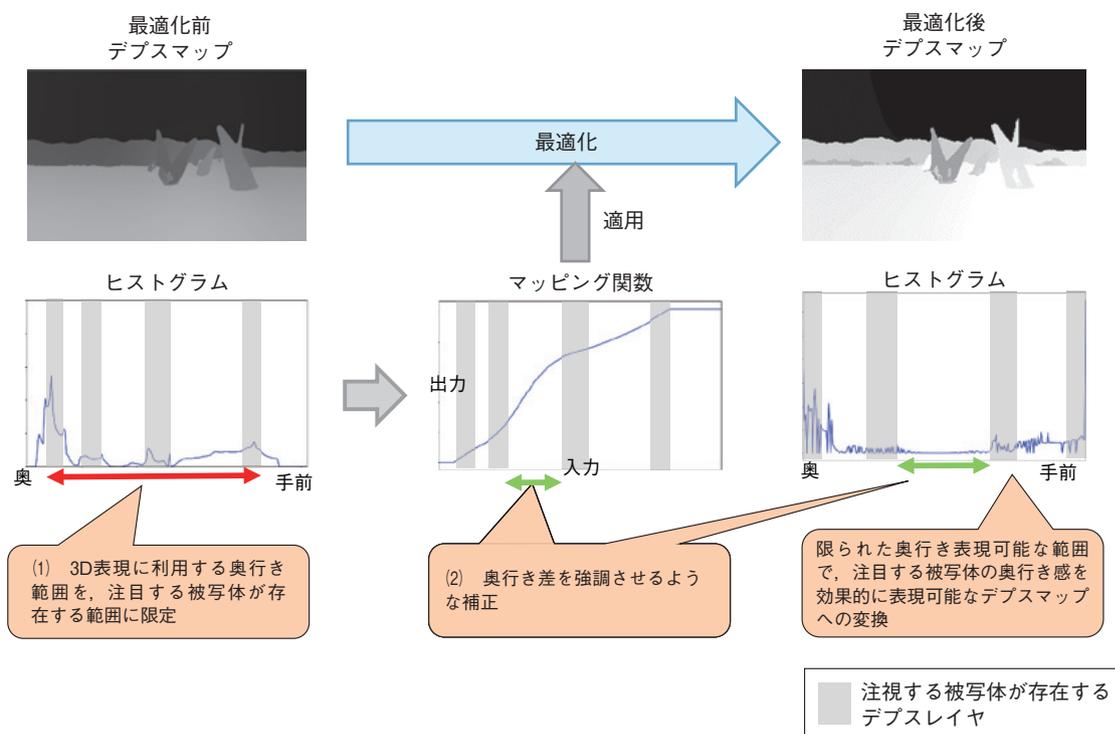


図3 デプスマップ最適化技術の概要

実際よりも大きくし、そのデプスレイヤにおける奥行表現がもっとも強調されるように、各デプスレイヤの幅や位置関係を補正します。その結果、デプス値の横軸を入力・縦軸を出力とするマッピング関数において、もっとも注目したい被写体が存在するデプスレイヤにおける傾きが大きくなったマッピング関数が生成されることとなります。

上記2点のデプスマップ補正を行うためのマッピング関数（非線形なデプスマップ変換関数）を導出し、デプスマップ画像に適用することにより、最適化されたデプスマップ画像を生成します。

市中の各種2Dアーカイブ映像には、さまざまなカメラワークやシーンチェンジが含ま

れるため、デプスマップの内容が時間の経過とともに大きく変化する傾向にあります。そのようなケースでも、効果的な3D表現を可能とするデプスマップの最適化を行えるよう、今後もさまざまなコンテンツへの適用評価を行い、デプスマップ最適化の品質向上を図っていきたくと考えています。

HiddenStereo 3D 映像生成システムへの活用

現在、私たちは前述した技術により、生成・最適化したデプスマップを用いてナチュラルな3D視聴を実現するHiddenStereo 3D映像生成システム（図4）の開発を推進しています。

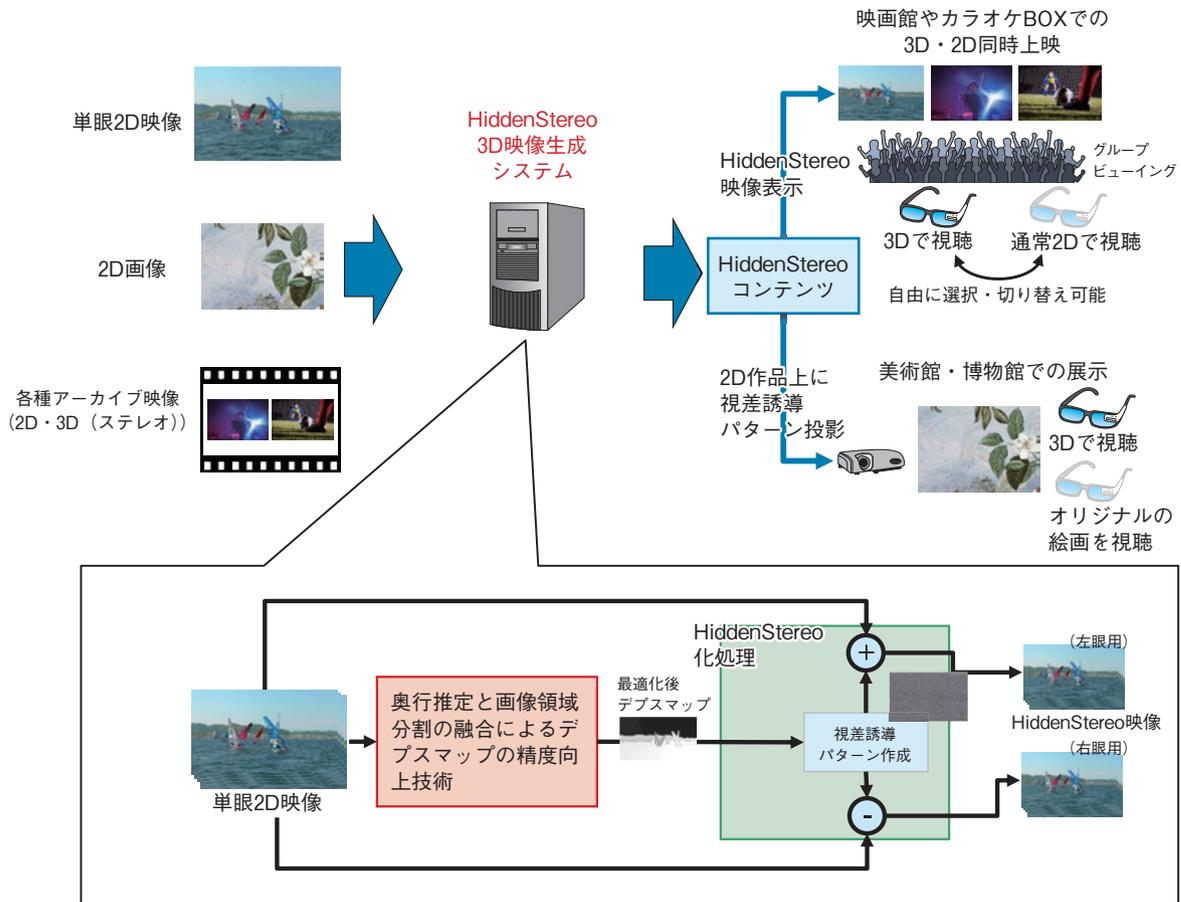


図4 HiddenStereo 3D映像生成システムの概要

HiddenStereoとは、NTTコミュニケーション科学基礎研究所の研究成果である、「3Dメガネをかけない視聴者には2D映像がクリアに見え、メガネをかけた視聴者には3D映像が見えるステレオ映像の生成技術」⁽³⁾です。本技術では、元画像とデプスマップから、人間に奥行き情報を与える働きをする視差誘導パターンを生成し、元画像に加算・減算することで、左目用・右目用画像を生成することができます。左右画像どうしを足し算すると視差誘導パターンが打ち消されて完全に元の画像に戻るため、3Dメガネをかけない視聴者はクリアな2D画像を見ることができ、メガネをかけた視聴者には、視差誘導パターンの効果でその画像に奥行きがついているように見えます。

本技術を活用することにより、1つの表示コンテンツに対して、その場にいる視聴者1人ひとりが楽しみ方を自由に選択できるという「人にやさしい3D表示」を実現可能です。また、2D・3D同時表示を行うために特殊な機器は不要であるため、既存の3D表示環境をそのまま活用でき、また通常の3D映像上映の際に必要な、3D上映と2D上映を別の時間帯とすることや、3D上映会場と2D上映会場を別に設定する必要がなくなり、3D映像上映における運用コストを低減することも可能となります。

HiddenStereoに対して、「奥行き推定と画像領域分割の融合によるデプスマップの精度向上技術」を活用することにより、新たにステレオカメラでの撮影や3DCG編集により制作したものだけでなく、過去に制作された2Dコンテンツの3Dコンテンツへの変換が可能となります。

今後の展開

奥行き推定と画像領域分割の融合によるデプスマップの精度向上技術と、本技術を用いて単眼2D映像からナチュラルな3D視聴を実現するHiddenStereo映像を生成するシステムを紹介しました。今後は、各技術のさらなる高速化や品質向上に向けた方式検討を進めるとともに、実証実験等を通じた評価を行い、ナチュラルなインタラクションを活用したビジネスの実現に貢献していきます。

参考文献

- (1) K.He, J.Sun, and X.Tang: "Guided Image Filtering," ECCV 2010, Heraklion, Greece, Sept. 2010.
- (2) G. Pestschnigg, R. Szeliski, M. Agrawala, M. Cohen, H. Hoppe, and K. Toyama: "Digital photography with Flash and no-Flash image pairs," ACM Transactions on Graphics, Vol. 23, No. 3, pp. 664-672, August 2004.
- (3) T. Fukiage, T. Kawabe, and S. Nishida: "Hiding of phase-based stereo disparity for ghost-free viewing without glasses," ACM Transactions on Graphics, Vol.36, No.4, pp.1-17, July 2017.



(左から) 小野 正人 / 菊地 由実 /
佐野 卓 / 深津 真二

より品質の高い3Dコンテンツの制作を実現するため、奥行き推定と画像領域分割の融合によるデプスマップの精度向上技術の完成度を高めるべく、研究開発を推進していきます。

◆問い合わせ先

NTTサービスエボリューション研究所
イノベティブサービス研究プロジェクト
TEL 046-859-2201
E-mail ev-journal-pb-ml@hco.ntt.co.jp