

# 透明視を利用した複合現実感環境下における遮蔽矛盾解消手法の提案

吹上大樹<sup>†1</sup> 大石岳史<sup>†1</sup> 池内克史<sup>†1</sup>

複合現実感 (MR) は、実世界に仮想物体 (CG) を重畳して表示することにより、高い臨場感を得ることができる技術である。しかしながら、任意の屋外環境において、その風景に CG を重ねると、しばしば CG と現実の物体との間で遮蔽関係に矛盾が生じるという問題がある。例えば、本来は前景であるはずの樹木が、重畳した CG に遮蔽されるというような状況は屋外では頻繁に生じ、MR の臨場感を著しく損なってしまう。この問題を解決するには、現実のシーンから前景となる部分の輪郭を正確に切り出す処理が必要となる。しかし、樹木や茂み等の複雑な形状を実時間で切り抜くことは計算量の点から実現が難しいというのが実状である。そこで本研究では、正確な前景情報を必要としない、新たな遮蔽矛盾解消手法を提案する。具体的には、心理物理実験によって測定したヒトの透明視知覚特性に基づいて、CG と現実風景とのブレンディングを行い、これによって CG が前景領域の奥に透けて見えるような描画を可能にした。

## Reduction of contradictory partial occlusion in Mixed Reality by using characteristics of transparency perception

TAIKI FUKIAGE<sup>†1</sup> TAKESHI OISHI<sup>†1</sup>  
KATSUSHI IKEUCHI<sup>†1</sup>

Mixed Reality (MR) is a technique that enables the merging of the real and virtual worlds by rendering virtual objects on a real scene in real time. One of the challenges in mixed reality (MR) applications is handling contradictory occlusions between real and virtual objects. The previous studies have tried to solve the occlusion problem by extracting the foreground region from the real image. However, real-time occlusion handling is still difficult since it takes too much computational cost to precisely segment foreground regions in a complex scene. In this study, therefore, we proposed an alternative solution to the occlusion problem that does not require precise foreground-background segmentation. In our method, a virtual object is blended with a real scene so that the virtual object can be perceived as being behind the foreground region. For this purpose, we first investigated characteristics of human transparency perception in a psychophysical experiment. Then we made a blending algorithm applicable to real scenes based on the results of the experiment.

### 1. はじめに

複合現実感 (MR) は、位置等の整合性を保ちながら、実時間で現実風景に仮想物体 (CG) を重畳する技術である。MR 環境下では、頭部装着型ディスプレイ等の装置を使うことで任意の方向を眺めることができ、非常にインタラクティブ性の高い臨場感のある体験が可能になるため、次世代の映像表現技術として現在盛んに研究が行われている。

しかし、CG と現実との完全な融合を果たすには現状ではまだ解決すべき課題が多く存在しており、その一つに遮蔽問題がある。遮蔽問題とは、座標上では奥に配置されている CG が、より手前に存在する現実物体を遮蔽してしまう問題である。これまでに、遮蔽問題を解決するために様々な手法が考案されており、例えばステレオ視や visual hull [1] 等の手法により前景を抽出する方法がある [2][3][4][5][6]。しかしこれらの手法は実時間動作が難しかったり、使用環境が予め定められていたりして、屋外の任意環境での使用は想定されていない。これに対し、[7][8] は単一カメラを用いて実時間で前景を抽出する手法を提案しており、[9][10]

はこれをさらに発展させ、屋外環境における動的な物体が生じる遮蔽問題を解決する手法を開発した。

このように、遮蔽問題を解決するため多くの研究が行われているが、任意の屋外環境において生じる遮蔽問題を完全に解決できる手法は現状では存在しない。屋外には樹木や茂み等の複雑な遮蔽物が多く存在し、これらを正確に切り出すには多大な計算量が必要となり、実時間処理が難しくなるためである。そこで本研究では、セグメンテーションの手法を改善するのではなく、祖末なセグメンテーション手法でも、自然に CG が遮蔽物の奥に見えるような CG と現実風景のブレンディング手法を提案する。本研究の最大の特色は、このブレンディングを実現するに当たり、ヒトの透明視知覚の特性を利用したことにある。

### 2. 透明視知覚特性の実験測定

#### 2.1 透明視知覚の x-junction モデル

ヒトの視覚系は非常に単純な 2 次元の絵を呈示するだけで、異なる奥行き位置に 2 つの半透明な面が存在しているような知覚を生じ得る (図 1 A と B)。過去の研究によれば、2 次元画像から重なり合った面を復元するとき、視覚系は非常に単純なヒューリスティクスを利用していると考えら

<sup>†1</sup> 東京大学  
University of Tokyo

れている[11][12][13]. それは2つのエッジが交差する「x-junction」周りの明るさのパターンに基づいており、例えば最も明るい面からより暗い面に線を引いていき、それがCの字を形成するようなら、視覚系は常に一方の面が半透明で、もう一方より手前にあるような透明視知覚を生じる(図1A, unique transparency). もし同じように引いた線がZの字を形成するなら、どちらの面が手前に見えるかは確定せず、ある時は一方が手前に見えていたものが、しばらくすると奥に見えたりする(図1B, bistable transparency). もし線が十字に交差したなら、どちらの面も透明には見えず、2次元のパターンがそのまま2次元の絵として知覚される(図1C, no transparency). こうしたヒューリスティクスは大雑把に物理的な光学的性質を反映していると考えられている[11][14].

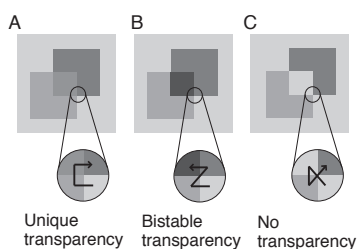


図1 X-junctionモデルによる透明視知覚の分類

Figure 1 Transparency perception classified based on patterns around an x-junction.

したがって、理想的には unique transparency を利用してCGが常に現実の前景の奥に見えるようなブレンディング方法をデザインできれば良いことになる。しかし、そのような状況を作るには、前景となる部分の輪郭の内側と外側でブレンディングの式を切り替えなければならない。そのためには正確な前景の情報が必要となるため、unique transparency を利用することは本研究の目的とそぐわない。

よって本研究では、bistable transparency を利用してブレンディング方法をデザインすることにした。後述するように bistable transparency は同一のブレンディングの式で実現できるため、前景の輪郭の複雑さに影響されない点で優れている。ただし、前述の通り bistable transparency では、常に一方の面が奥に見えるという保証がない。しかし過去の研究によって、bistable transparency においても、「一方の面が奥に見える確率」が x-junction を構成する各エッジのコントラストに依存して変動することが分かっている[14][16]. もし、x-junction 周辺のコントラストを引数として、「一方の面が奥に見える確率」が得られるような透明視のモデルを作ることができれば、bistable transparency を利用してCGが前景の奥に見えるブレンディング方法をデザイン可能になるだろう。

## 2.2 透明視知覚特性の測定とモデル化

X-junction 周りの明るさを引数として、ある面がもう一方の面よりも奥に見える確率が得られる透明視知覚のモデルを作るため、本研究では心理物理学的な実験手法を用いて測定を行った。過去の研究でも、bistable transparency 図形を用いて透明視知覚特性を測定したものがいくつか見られるが[14][15][16], 任意の輝度値を引数とした知覚モデルを作成するにはデータが不十分であった。したがって本実験では、あらゆるパターンを網羅できるように非常に多くの実験刺激を用いて測定を行った。

### 2.2.1 実験方法

**実験参加者:** 実験目的を知らない22-25歳の8名(男性7名, 女性1名)が実験に参加した。

**実験装置:** 刺激は暗室中のCRTモニタ(Sony Trinitron Multiscan CPD-17SF9, 17インチ, 1024 x 768 pixels, リフレッシュレート75Hz, 平均輝度44.6 cd/m<sup>2</sup>)上に呈示した。参加者は顎台で頭を固定し両眼で刺激を観察した。観察距離は86cmであった。

**実験刺激:** 刺激は1つの円(直径4deg)と1つの縦長の長方形(4deg x 8deg)からなっており、円のちょうど右半分が長方形と重なるように配置されていた。したがって、刺激は背景部分(B), 長方形(R), 円(D), それらが重なった部分(O)の4つの領域をもっていた。これら4つの領域の輝度を適切に設定することで、bistable transparency 刺激を作ることができる。図3に、可能な全ての bistable transparency 刺激のパターンを示した。これら4つのタイプの各領域の輝度を様々に変化させることで、計438パターンの刺激を用意した。

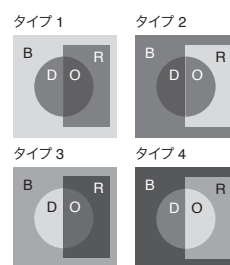


図2 実験刺激の例

Figure 2 Examples of the experimental stimuli.

**実験手順:** 各試行の流れは、まず0.75秒間刺激を呈示し、直後に99%コントラストのランダムドットパターンを呈示した。その間に実験参加者は「円が長方形よりも奥に見えるか手前に見たか」を2肢強制選択で回答した。ランダムドットパターンは特定の明るさへの順応を防ぐために用いた。実験参加者がキー押しで回答するとすぐに次の試行が開始された。1つのセッションで、全438種類の刺激がランダムな順でテストされた。各実験参加者は全部で6セッション分実験を行い、1つの刺激に対し8人で48回分

の回答が集められた。これらの回答から、「円が長方形よりも奥に見えた確率」を計算した。

### 2.2.2 結果のモデル化

実験で得られた「円が長方形よりも奥に見えた確率」は、刺激によって大きく異なっていた。この確率の変動を説明可能なモデルはあるだろうか？[16]は、本研究と同様の実験をより少ない刺激(タイプ2とタイプ3のみから20種類)で行い、その結果に基づいて「円が長方形よりも奥に見える確率」が明るさのコントラストの差分 $(|D \cdot B| + |D \cdot O|) - (|R \cdot B| + |R \cdot O|)$ に従って増加するというモデルを提案した。しかしながら、本研究の実験結果はこのモデルで説明できない部分が多かった。そこで、本研究では、より柔軟な以下のモデルを提案する。

$$X = w \frac{c(D,O)}{c(D,O) + c(R,O)} + (1-w) \frac{c(D,B)}{c(D,B) + c(R,B)} \quad (1)$$

「円が長方形よりも奥に見える確率」は式(1)で得られる $X$ に従って増加する。式(1)の $w$ は重み付け変数で、 $B, R, D, O$ は実験刺激の各領域の輝度値を表す(図2参照)。 $c(D,O)$ は領域 $D$ と $O$ の間の Michelson contrast を表し、次のように計算される。

$$c(D,O) = \frac{|D - O|}{D + O} \quad (2)$$

Michelson contrast は感覚量に基づいたコントラストの定義としてよく用いられる[17][18]。

式(1)に表されるように、このモデルは「円が奥に見える確率」が2つの要素の重み付け和に従って増加すると想定している。1つ目の要素は、領域 $D$ と $O$ の間のコントラストが $R \cdot O$ 間のコントラストと比べてどれくらい大きいかを表し、2つ目の要素は、 $D \cdot B$ 間のコントラストが $R \cdot B$ 間のコントラストに比べてどれくらい大きいかを表す。式(1)で計算した $X$ を横軸にとって「円が奥に見えた確率」をプロットし、重み付け変数 $w$ を自由変数として以下のシグモイド関数 $S$ をフィットしたものを図3に示す。

$$S(x; a, b) = \frac{1}{1 + \exp\left(-\frac{x - a}{b}\right)} \quad (3)$$

図2の刺激タイプによって若干結果が異なっていたため、タイプ毎に異なるシグモイド関数をフィットした。ベストフィットパラメータは、 $w = 0.93$ 、タイプ1のシグモイド関数は $(a, b) = (0.41, 0.11)$ 、タイプ2は $(a, b) = (0.57, 0.14)$ 、タイプ3は $(a, b) = (0.55, 0.13)$ 、タイプ4は $(a, b) = (0.52, 0.12)$ であった。重み付け変数が1に近い値となったのは、知覚

される奥行き順序が、領域 $O$ (2つの面の重なり部分)と隣接する領域とのコントラストの値に大きく依存することを示唆している。

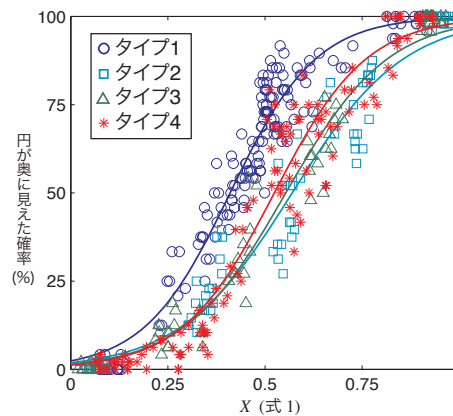


図3 提案モデル(式1)に基づいてプロットした「円が奥に見えた確率」

Figure 3 The percentage of disk-behind perception plotted according to the proposed model (Eq. 1).

### 3. 透明視知覚モデルに基づいたブレンディング手法

このセクションでは、実験で得られた bistable transparency 刺激に対する透明視知覚モデルに基づき、CG が現実風景中の前景よりも奥に見えるようなブレンディング手法を提案する。以降の説明の理解を助けるため、以下の図4に、これまで実験刺激として用いてきた単純図形が、実際のMRシーンにおけるブレンディング結果とどのように対応付けられるのかを示した。以後はこの図に基づき、実験刺激の長方形( $R$ )は現実風景画像中の前景部分(遮蔽物)、背景( $B$ )は現実風景画像の背景部分、円( $D$ )は背景とCGのブレンディング結果、重なり部分( $O$ )は前景とCGのブレンディング結果を表すものとする。

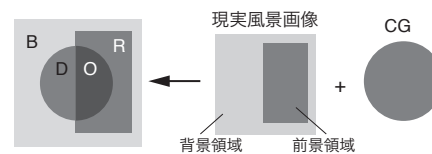


図4 MRシーンと実験刺激の対応付け

Figure 4 How an actual MR scene corresponds to the experimental stimuli we used.

CG を現実風景画像とブレンディングする際に bistable transparency の状態を生むような計算式はいくつか考えられるが、本研究では以下の2つの式を用いた。以後、式(4)によるブレンディングを「Multiplicative blending」、式(5)によるブレンディングを「Inversed-multiplicative

blending」と呼ぶ。

$$I_M(I_r, I_v) = I_r I_v, \quad (4)$$

$$I_I(I_r, I_v) = 1 - (1 - I_r)(1 - I_v), \quad (5)$$

ここで  $I_M$ ,  $I_I$  は、それぞれのブレンディングにより得られた画像の画素値を表し、 $I_r$ ,  $I_v$  はそれぞれ現実風景画像の画素値と CG の画素値を表す。なお、画素値は全て 0~1 の範囲にスケールされているものとする。これら 2 つのブレンディング結果は、いずれも必ず CG と前景の 2 つの面を **bistable transparency** の状態にすることができる。例えば、現実風景中の前景部分が背景部分よりも暗い画像に対して **multiplicative blending** で CG をブレンディングすると、図 2 のタイプ 1 のパターンができる。同じ画像に **inversed-multiplicative blending** を適用すると、図 2 のタイプ 3 のパターンができる。逆に、現実風景中の前景部分が背景部分よりも明るい画像に対して **multiplicative blending** で CG をブレンディングすると、図 2 のタイプ 2 のパターンができる。これに **inversed-multiplicative blending** を適用すると図 2 のタイプ 4 のパターンができる。このように、これら 2 つのブレンディング方法は、2 種類の異なる現実風景のパターンとの組み合わせにより、全てのタイプの **bistable transparency** パターンを網羅することができる。

しかしながら、このままでは与えられた CG と現実風景画像に対し、2 つのブレンディング結果から 1 つを選択することしかできず、知覚モデルに基づく調整を行う余地がほとんどない。そこで、以下のように式(4)と(5)の計算式に新たな可変パラメータ  $\alpha$  を導入することで、調整の自由度を与えた。

$$I_M'(I_r, I_v, \alpha) = \alpha I_M(I_r, I_v) + (1 - \alpha) I_r \quad (6)$$

$$I_I'(I_r, I_v, \alpha) = \alpha I_I(I_r, I_v) + (1 - \alpha) I_r \quad (7)$$

これらの式では、式(4), (5)で表されるそれぞれのブレンディング結果を再度現実風景画像とブレンディングしており、パラメータ  $\alpha$  はその透明度を表している。本研究の実験で、**bistable transparency** 刺激における 2 つの面のうち、どちらが奥に見えやすいかは、2 つの面の重なった領域に対するコントラスト (D-O v.s. R-O) に大きく依存することが分かったため、理論上はこのパラメータ  $\alpha$  を減少させることで、CG が奥に見える確率を単調に増加させることができる。

式(6), (7)を用いると、CG の画素値  $I_v$  と現実風景画像

の画素値 (背景部分を  $I_b$ , 前景部分を  $I_f$  とする) が与えられたとき、図 4 のブレンディング結果の各領域 (B, R, D, O) は以下のように表すことができる。

$$\begin{cases} B = I_b \\ R = I_f \\ D = I_M'(I_b, I_v, \alpha) \\ O = I_M'(I_f, I_v, \alpha) \end{cases} \quad (8)$$

$$\begin{cases} B = I_b \\ R = I_f \\ D = I_I'(I_b, I_v, \alpha) \\ O = I_I'(I_f, I_v, \alpha) \end{cases} \quad (9)$$

式 (8) は **multiplicative blending**, 式 (9) は **inversed-multiplicative blending** によるブレンディング結果を表す。これらの値を式(1)に代入し、さらに得られた  $X$  を図 3 に表されたシグモイド関数に代入することで、CG が奥に見える確率がパラメータ  $\alpha$  の関数として得られる。すなわち、任意の  $I_b$ ,  $I_f$ ,  $I_r$  が与えられたとき、知覚モデルに従って CG が奥に見えやすくなるように  $\alpha$  の値を一意に定めることができる。

図 5 に、様々な  $I_b$ ,  $I_f$  の組み合わせについて、50%以上の確率で CG が前景の奥に見えるために必要な  $\alpha$  の上限値をシミュレーションによって求めてプロットしたものを示した。図 5 A は **multiplicative blending** によって得られた結果を、図 5 B は **inversed-multiplicative blending** によって得られた結果を表している。なお、このシミュレーションにおいては、**multiplicative blending** では  $I_v$  を 0 に、**inversed-multiplicative blending** では  $I_v$  を 1 に設定しているが、他の値でシミュレーションを行っても定性的な結果は変わらなかった。

次に、**multiplicative blending** と **inversed-multiplicative blending** のどちらを使用すべきかについて考える。そのための重要な判断基準として、CG の視認性が挙げられる。つまり、高い視認性を維持しつつ、CG が奥に見えやすいブレンディング結果が得られる方を選択すればよい。基本的に、 $\alpha$  値が大きい程視認性は高くなるが、常にそうであるとは限らない。例えば、**multiplicative blending** においては、現実風景画像の背景が暗ければ、 $\alpha$  が 1 であっても CG の視認性は下がる。逆に **inversed-multiplicative blending** では、背景が明るければ  $\alpha$  が 1 でも CG は見え辛くなってしまふ。このため、視認性を評価するため、以下の式で定義される指標を用いた。

$$V_M = \alpha I_b \quad (10)$$

$$V_I = \alpha(1 - I_b) \quad (11)$$

ここで、 $V_M$  は **multiplicative blending** の視認性、 $V_I$  は

inversed-multiplicative blending の視認性を表す. 図5の縦軸をこれらの指標に差し替えてプロットし直すと, 図6のようになる. 図6A, B から分かる通り, 現実風景画像の前景部分が背景部分よりも暗いときは, multiplicative blending の方が高い視認性を示し, 逆に前景が背景よりも明るいときは inversed-multiplicative blending の方が高い視認性を示している. したがって, 現実風景画像の前景, 背景のどちらが明るいかを基準として, ブレンディングの計算式を選択するべきであると言える.

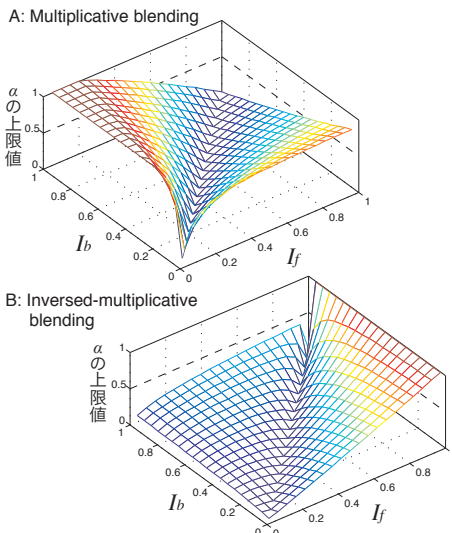


図5 50%以上の確率でCGが奥に見える $\alpha$ の上限値  
Figure 5 Upper limits of  $\alpha$  with which an observer perceives the CG as behind in more than 50% chances plotted as a function of  $I_b$  and  $I_f$ .

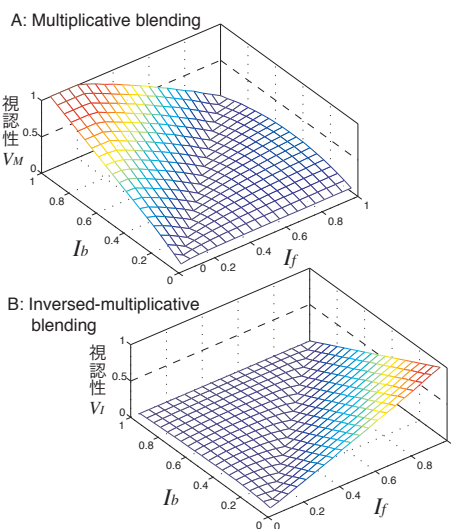


図6 図5の $\alpha$ 値におけるCGの視認性  
Figure6 Visibility of the CG blended with  $\alpha$  values in Fig. 5.

まとめると, 本研究で提案するブレンディング方法は以下のようなになる.

1. **入力:** 現実風景画像中の前景部分の明るさ ( $I_f$ ), 背景部分の明るさ ( $I_b$ ), CGの明るさ ( $I_v$ ) あ
2. **ブレンディング方法の選択:** もし  $I_f < I_b$  なら multiplicative blending (式6) を, もし  $I_f > I_b$  なら inversed-multiplicative blending (式7) を使用する
3. **ブレンディングパラメータ $\alpha$ の決定:** まず, ブレンディングによって生じる x-junction (CGのエッジと前景領域のエッジの交差する所) 周りの4つの領域 (B, R, D, O) の明るさを式(8)もしくは式(9)によって計算する. 次に式(1)にこれらの値を代入して  $X$  を求め, これをさらに実験データにフィットしたシグモイド関数に代入し, CGが前景の奥に見える確率を $\alpha$ の関数として得る (Multiplicative blending の場合はタイプ1のデータにフィットさせたシグモイド関数を用い, inversed-multiplicative blending の場合はタイプ4のデータにフィットさせた関数を用いる). 最後に, その確率が50%以上になるような最大の $\alpha$ 値を求め, これをブレンディングに使用する.

#### 4. 提案手法の実装と実験

このセクションでは, セクション3で提案したブレンディング方法を実際のアプリケーションとして実装するアルゴリズムと, その実験結果について述べる.

##### 4.1 ブレンディングアルゴリズムの実装

提案手法では, 現実風景画像中の前景部分と背景部分の明るさを知るため, 現実風景画像の前景領域を示す確率分布画像が必要となる. しかし, この確率分布は正確である必要はない. 確率分布画像は深度マップやセグメンテーション, オプティカルフロー等の様々な方法で得ることができるが, ここではその手法については問わず, 既に確率分布画像が得られているという前提の元に話を進めていくこととする.

基本的に, 本研究で提案したブレンディングは現実風景画像中の前景部分と背景部分が単色で構成されている場合に最良の結果を生むが, 現実にはそのような場面はほとんど存在しない. したがって, 本研究ではこの制約を克服するため, 画素単位でブレンディング方法を決めるというアルゴリズムによって実装を行った. すなわち, 提案手法では, 画像の各画素をスキャンしていき, 画素毎にその画素を含むある大きさの窓内で背景, 前景およびCGの明るさを計算し, ブレンディングの計算式とパラメータ $\alpha$ を決定する. 隣り合う画素の窓同士はほとんどの画素を共有するため, ブレンディングパラメータは滑らかに変化する. 仮に隣り合う画素間でブレンディングの式が異なっても, そのような場合は前景と背景の明るさが等しくなっているため, CGがほとんど完全に透明になり, ブレンディングによって意図しないエッジが生じることはない.

以下に, ブレンディングアルゴリズムについて詳しく述

べる。スキヤニング処理中における現在の画素を $(x,y)$ とし、 $P_r$ ,  $P_v$ ,  $P_m$ をそれぞれ現実風景画像、CG 画像、前景確率分布画像を表すものとする。まず、 $P_r(x,y)$ ,  $P_v(x,y)$ ,  $P_m(x,y)$ の各画素について、その画素を中心としたあるサイズの正方形窓内の画素を調べ、CG の画素値の平均値  $I_v$ 、背景領域の画素値の平均値  $I_b$ 、前景領域の画素値の平均値  $I_f$ を以下の式に従って求める。

$$I_v = \frac{1}{\sum_{(p,q) \in W} A_v(p,q)} \sum_{(p,q) \in W} P_v(p,q) A_v(p,q) \quad (12)$$

$$I_b = \frac{1}{\sum_{(p,q) \in W} \{1 - P_m(p,q)\}} \sum_{(p,q) \in W} P_r(p,q) \{1 - P_m(p,q)\} \quad (13)$$

$$I_f = \frac{1}{\sum_{(p,q) \in W} P_m(p,q)} \sum_{(p,q) \in W} P_r(p,q) P_m(p,q) \quad (14)$$

ここで、 $W$  は平均窓内の画素の集合を表している。また、 $A_v$ は CG 画像のアルファチャンネル配列を表す。 $A_v$ は CG が描画されたオフスクリーンバッファ内で、CG が存在する画素のみから平均値を求めるために使用している。これら3つの値に従い、前セクションで述べた方法により、画素 $(x,y)$ におけるブレンディング計算式とブレンディングパラメータ  $\alpha$  が決定される。この  $\alpha$  を用いることで、画素 $(x,y)$ におけるブレンディング結果は以下ようになる。

$$P_{blend} = \begin{cases} \alpha P_v(x,y) P_r(x,y) + (1 - \alpha) P_r(x,y), & \text{if } I_f \leq I_b \\ \alpha [1 - \{1 - P_v(x,y)\} \{1 - P_r(x,y)\}] + (1 - \alpha) P_r(x,y), & \text{if } I_f > I_b \end{cases} \quad (15)$$

もしある画素を中心とした平均窓の全て、あるいはほとんどが前景領域内にあった場合、背景領域の平均画素値が得られないため、以上の方法ではブレンディングの式や  $\alpha$  を決定できない。そのような時は、窓内のある割合 ( $\tau$ ) 以上の部分が背景領域を含むようになるまで、平均窓のサイズを 1.5 倍ずつ拡張する。この拡張された窓は、 $I_b$  の計算のみに使用される。この処理により、隣接する画素間で平均窓サイズが異なる場合が生じうるが、これはブレンディング結果に意図しないエッジをもたらすことがある。これを防ぐために、最終的な背景領域の平均画素値 ( $I_b'$ ) は、以下のように全ての平均窓サイズにおいて計算された  $I_b$  を重み付け平均することで求める。

$$I_b' = \left( 1 - \sum_{k=0}^{n-1} \rho_k \right) I_{b:W_n} + \sum_{k=0}^{n-1} \rho_k I_{b:W_k} \quad (16)$$

ここで、 $W_n$  は初期サイズの平均窓 ( $W_0$ ) を 1.5<sup>n</sup> 倍した平均窓を表し、 $n$  は十分な大きさが得られるまでに平均窓が

拡張された回数を表す。また、 $I_{b:W_n}$  は  $W_n$  において計算された  $I_b$  を表す。重み付け  $\rho$  は、以下のように定義される。

$$\rho_k = \frac{r_k^2}{\tau \sum_{l=0}^{n-1} r_l} \quad (17)$$

ここで、 $r_k$  は  $W_n$  内において背景領域が占める割合で、

$$r_k = \frac{1}{N} \sum_{(p,q) \in W_k} \{1 - P_m(p,q)\} \quad (18)$$

によって得られる。この処理により、平均窓サイズが異なる画素間でブレンディング結果が滑らかになる。

これとは逆に、もし平均窓内の全ての画素が背景領域内であった場合、その部分では遮蔽問題は生じないため、ブレンディングを行う必要はなくなる。したがってその画素では、CG の画素値がそのままブレンディング結果画像に代入される。ただし、ブレンディングを行わない画素とブレンディングを行う画素の間で結果を滑らかにするため、以下の式を導入した。

$$P_{output}(x,y) = (1 - \lambda) P_{blend}(x,y) + \lambda P_v(x,y) \quad (18)$$

ここで  $P_{output}(x,y)$  は画素 $(x,y)$ における最終的なブレンディング結果を表す。 $\lambda$  はブレンディングを行う画素とそうでない画素を滑らかに切り替えるための重み付け関数で、

$$\lambda = S \left( \frac{\sum_{(p,q) \in W} (1 - P_m(p,q))}{N}; \text{threshold, slope} \right) \quad (19)$$

のように定義される。ここで  $N$  は平均窓内のピクセルの数を表し、 $S$  は式(3)で定義されるシグモイド関数である。

## 4.2 実験

### 4.2.1 実験環境

以上のように定義されたアルゴリズムをテストするため、前景領域が遮蔽問題を生じる様々な現実風景画像を使って実験を行った。画像解像度は 640x480 ピクセルだった。前景確率分布画像はペイントソフトを用いて手動で作成したが、単に CG を切り抜くための前景マスクとして用いるには不十分な精度となるよう、意図的に大雑把な画像にした。実験には一般的な PC (OS: Windows 7, CPU: Corei7 2.93 GHz, RAM: 8GB, GPU: nVIDIA GTX 550Ti 1024MB) を用いた。平均窓の初期値  $W_0$  は 60x60 ピクセルとした。平均窓サイズを決定するための背景領域の割合の下限值 ( $\tau$ ) は 0.1 とした。式(19)のシグモイド関数のパラメータ (threshold, slope) は、それぞれ 0.9 と 0.05 とした。

ブレンディングアルゴリズムは全てプログラマブルシェーダ (GLSL) 内に記述することで高速化を図った。さらに、実時間動作を実現するため、 $I_b$ ,  $I_f$ ,  $I_v$  を計算する際は平均窓内の画素を 6 画素毎にサンプリングして平均値を求めた。実際の FPS は前景領域と CG が重なる領域の画素数

に大きく依存するが、以上のように実装することで、多くの場合 25FPS 以上で動作した。CG と前景領域が画面全体を占める場合は、8 FPS ほどに落ち込んだが、これは平均窓内の画素のサンプリングレートを適応的に変化させることで改善可能であると考えられる。

#### 4.2.2 実験結果

提案手法によるブレンディング結果の画像を図7の一番左の列に示した。結果を比較するため、その右隣の列には、提案手法で用いた前景確率分布画像をアルファチャンネルマスクとして用い、ファルファブレンディングを行った結果を示した。左から3番目の列はCGをブレンディングする前の現実風景画像、一番右側の列はブレンディングに用いた前景確率分布画像である。図7A, B は前景確率分布が比較的正確な場合を想定しており、図5C は前景確率分布がやや曖昧で、100%に満たない場合を想定して確率分布画像を作成した。結果から、前景領域を示す情報が非常に曖昧であっても、提案手法によるブレンディングによって、CGの奥行き位置が正しく知覚されやすいMRシーンを合成できることがわかる。特に、提案手法は前景の奥にCGが透けて見えるようなブレンディング結果を生成するため、輪郭の形状の複雑さが結果に大きく影響することがない。一方、単純なアルファブレンディングによる結果では、前景が複雑な形状をもっていた場合に、その輪郭周辺でCGが見えなくなり、葉の隙間等、実際はCGが見えているはずの部分に不自然な空白を生じる場合がある。また、前景と判定された領域が実際よりも狭かったり、前景確率が曖昧で、100%に満たなかったりした場合(図7C)は、単純なアルファブレンディングでは遮蔽問題を生じてしまう危険が大きいのにに対し、提案手法では安定してCGが前景の奥に見える結果を生成できることが確かめられた。

#### 4.2.3 解決すべき課題

以上のような利点はあるものの、現段階では提案手法にはまだ解決すべき課題もある。まず、提案したブレンディング方法では、前景領域と背景領域の明るさがほとんど同じであった場合、 $\alpha$ 値が0に近くなり、CGが不可視になってしまう。そのような時は、ブレンディング結果は単純なアルファブレンディングと大差ないものとなる。この点を改善するには、本研究で利用した「x-junction 周りの明るさ分布」とは異なる奥行き手がかりを利用して、視認性を保ったままより正しい奥行き知覚が得られるよう、ブレンディング方法を改良していく必要がある。

また、本研究では、単純な刺激を用いた心理物理実験結果に基づいたモデルを元にブレンディングパラメータの調整を行っているが、現実風景には非常に多くの奥行き手がかりが存在するため、この調整がどれほど妥当であるかは不明である。今後、現実風景画像との合成シーン等を刺激として用いた心理実験を行い、さらに調整を行う必要がある。

最後に、本研究で提案したブレンディングアルゴリズムの実装方法では、静止画に対するブレンディングのみを想定していたため、映像に対してこれを適用するとブレンディング結果が時間的に滑らかに繋がらない可能性が高い。MRアプリケーションは通常、実時間で取得した映像に対してCGを重ねるため、このままでは実用的といえない。しかし、この問題は $I_b$ ,  $I_f$ を計算する平均窓を時間方向に拡張することで容易に解決できると考えられる。精度は悪いが実時間で前景確率分布が得られるような手法と動画に対応した提案手法を組み合わせ、実用的なMRアプリケーションを実装することが将来的な課題である。

## 5. 結論

本研究では、心理物理実験によって測定した透明視知覚特性のモデル化を行い、それに基づいてCGが現実の遮蔽物の奥に知覚されるような、CGと現実風景のブレンディング手法を提案した。提案手法では、曖昧な前景確率分布を与えるだけで、複雑な輪郭をもった遮蔽物が存在するシーンにおいても、効率的に遮蔽矛盾を抑えることが可能となることが実験により確かめられた。このブレンディング手法と低コストの前景判定アルゴリズムを組み合わせることで、任意のシーンでも実時間で遮蔽問題を処理可能なMRシステムが実現できると考えられる。

## 参考文献

- 1) A. Laurentini. The visual hull concept for silhouette-based image understanding. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16: 150-162, 1994.
- 2) M. Kanbara and N. Yokoya. Geometric and photometric registration for real-time augmented reality. In *ISMAR*, pages 279-280, 2002.
- 3) H. Kim, S. J. Yang, and K. Sohn. 3d reconstruction of stereo images for interaction between real and virtual worlds. In *ISMAR*, pages 169-177, 2003.
- 4) J. Allard, C. Menier, B. Raffin, E. Boyer, and F. Faure. Grimage: Markerless 3d interactions. In *ACM SIGGRAPH Emerging Technologies*, 2007.
- 5) J. M. Hasenfratz, M. Lapierre, F. Sillion. A real-time system for full body interaction with virtual worlds. *Eurographics In Symposium on Virtual Environments*, pages 147-156, 2004.
- 6) A. Ladikos and N. Navab. Real-time 3d reconstruction for occlusion-aware interactions in mixed reality. In *ISVC (1)*, pages 480-489, 2009.
- 7) A. Criminisi, G. Cross, A. Blake, and V. Kolmogorov. Bilayer segmentation of live video. In *CVPR (1)*, pages 53-60, 2006.
- 8) J. Sun, W. Zhang, X. Tang, and H. Y. Shum. Background cut. In *ECCV (2)*, pages 628-641, 2006.
- 9) T. Kakuta, L. B. Vinh, R. Kawakami, T. Oishi, and K. Ikeuchi. Detection of moving objects and cast shadows using a spherical vision camera for outdoor mixed reality. In *VRST*, pages 219-222, 2008.
- 10) L. B. Vinh, T. Kakuta, R. Kawakami, T. Oishi, and K. Ikeuchi. Foreground and Shadow Occlusion Handling for Outdoor Augmented Reality. In *ISMAR*, pages 109-118, 2010.
- 11) E. H. Adelson and P. Anandan. Ordinal characteristics of transparency. In *AAAI-90 Workshop on Qualitative Vision*, 1990.
- 12) B. L. Anderson. A theory of illusory lightness and transparency in monocular and binocular images: the role of contour junctions. *Perception*, 26(4): 419-453, 1997.

- 13) J. Beck and R. Ivry. On the role of figural organization in perceptual transparency. *Perception & Psychophysics*, 44(6): 585-594, 1988.
- 14) A. Kitaoka. A new explanation of perceptual transparency connecting the X-junction contrast-polarity model with the luminance-based arithmetic model. *Japanese Psychological Research*, 47(3): 175-187, 2005.
- 15) J. Beck, K. Prazdny, and R. Ivry. The perception of transparency with achromatic colors. *Perception & Psychophysics*, 35(5): 407-422, 1984.
- 16) F. Delogu, G. Fedorov, M. O. Belardinelli, and C. van Leeuwen. Perceptual preferences in depth stratification of transparent layers: Photometric and non-photometric factors. *Journal of Vision*, 10(2): 1-13, 2010.
- 17) J. Rovamo, O. Luntinen, and R. Näsänen. Modelling the dependence of contrast sensitivity on grating area and spatial frequency. *Vision Research*, 33(18): 2773-2788, 1993.
- 18) M. Singh and B. L. Anderson. Toward a perceptual theory of transparency. *Psychological Review*, 109: 492-519, 2002.

Proposed blending method    Simple alpha blending

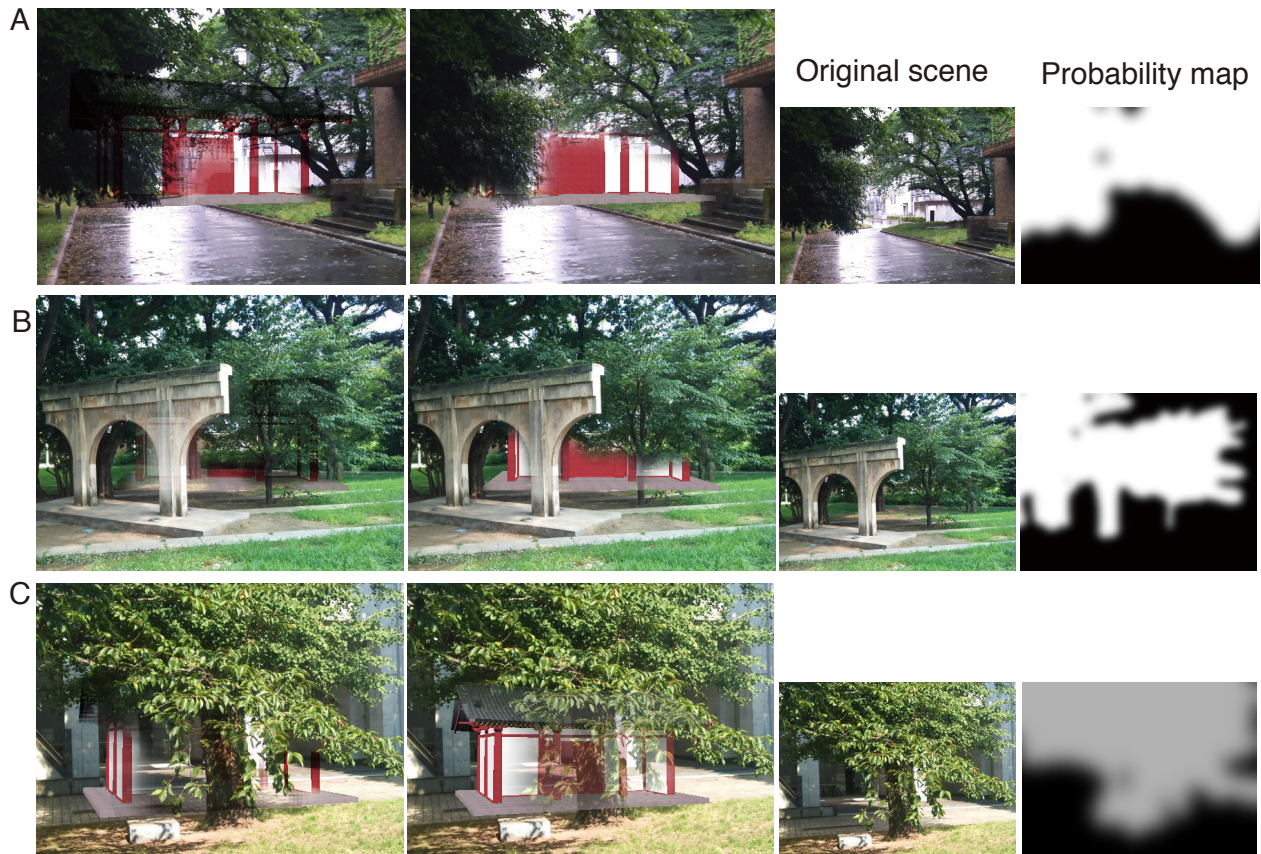


図7 1列目：提案手法によるブレンディング結果 2列目：既存手法（アルファブレンディング）による結果 3列目：ブレンディング前のシーン 4列目：ブレンディングに使用した前景確率分布画像

Figure7 1<sup>st</sup> column: results of our blending methods. 2<sup>nd</sup> column: comparison results obtained by simple alpha blending. 3<sup>rd</sup> column: original images of real scenes. 4<sup>th</sup> column: probability maps of foreground regions.