

Mixed Reality における現実物体の物理属性を考慮した シーン生成

岡部 聡一^{†1} 宮脇 健三郎^{†2}

概要： Mixed Reality は、現実を拡張する Augmented Reality や現実と異なる仮想空間を体験させる Virtual Reality と異なり、仮想と現実を融合させた表現を行うことができる。例えば CG モデルのリンゴを現実の机に置く動作をした場合、リンゴは現実の机との物理的な相互作用を起こした上で、上に乗り、そのリンゴをあらゆる角度から見る事が可能である。しかし、現状の Mixed Reality はあくまで現実空間の位置や物体の形状などの認識にとどまっており、仮想物体に対し与える影響は現実空間の物体との衝突などにとどまる。そこで本研究では仮想物体をより現実存在するかのように表現するための新たな手法を提案する。具体的には現実空間に存在する物体の認識とその構成素材などを把握することで硬度や摩擦力を算出し、反映させることでより現実的な表現を行う。

Generation of Mixed Reality Scene with Objects' Physical Characteristics

SOICHI OKABE^{†1} KENZABURO MIYAWAKI^{†2}

Abstract: Mixed Reality can express expressions that combine virtual and real world, and it is different from Augmented Reality and Virtual Reality. For example, in the Mixed Reality world, user can put a virtual CG apple on a desk of the real world. The virtual apple can interact with the real desk, and user can view the virtual apple for every direction. However, the Mixed Reality only uses position and shape information of real world objects. Therefore, we present a new Mixed Reality scene generator that can express objects' physical characteristics such as frictional, hardness and so on.

1. はじめに

近年、Augmented Reality や Virtual Reality の発展・拡張要素として Mixed Reality が注目されている。Mixed Reality は、現実を拡張する Augmented Reality や現実と異なる仮想空間を体験させる Virtual Reality と異なり、仮想と現実を融合させた表現を行うことができる例えば CG モデルのリンゴを現実の机に置く動作をした場合、リンゴは現実の机との物理的な相互作用を起こした上で、上に乗り、そのリンゴをあらゆる角度から見る事が可能である。また、現実の机の上に他の物体があり、その陰にリンゴを置くと、リンゴは陰になっている部分だけ表示されず、現実物体の近くにあるように見ることが出来る。しかし、現状の Mixed Reality はあくまで現実空間の位置や物体の形状などの認識にとどまっており、仮想物体に対し与える影響は現実空間の物体との衝突などにとどまる。そこで本研究では仮想物体をより現実存在するかのように表現するための新たな手法を提案する。

具体的には現実空間にある物体の硬度や摩擦力などを考慮することで、より現実存在すると感じることができ

るのではないかと考える。例えばスーパーボールを現実のソファの上に落とすシチュエーションを想定する。ソファに落ちたボールはクッションの影響を受けてあまり跳ねず、転がる時も布地との摩擦ですぐに止まると考えられる。これを仮想物体のボールを用いて既存の Mixed Reality で行うとソファの上に落ちた場合でも床に落ちた場合でも変わらず一定の跳ね方、転がり方となる。これはソファが柔らかく衝撃を吸収しやすく、布地との摩擦で動きが停止しやすいことを考慮されていないためである。故に本研究では現実空間に存在する物体の認識及びその構成素材などを把握することにより、硬度、衝突音、摩擦力を想定し、反映させることで仮想物体が現実存在するかのような表現を実現していくことを目指すものである。

2. 関連研究

Mixed Reality に関連する研究において、仮想物体がより現実的に見える表現を行う研究がいくつかある。例えば角田 [1]は Mixed Reality 技術を用いて文化財の復元を行っているが、復元する際に天空の光源分布を元に仮想物体の影画像をリアルタイムで生成することで、実空間に対して自然な陰影付けを可能とし、仮想物体の現実感向上を図っている。また、稲葉らの研究では実物体の 3 次元形状、アルベド値、光源環境を持つ実世界環境モデルから現在の見えを再現することによって、光源環境変化に頑健な特徴点マ

^{†1} 大阪工業大学大学院 情報科学研究科

Major in Information Science and Technology, Graduate School of Osaka Institute of Technology

^{†2} 大阪工業大学 情報科学部

Faculty of Information Science and Technology, Osaka Institute of Technology

ッチングを実現しており[2]、整合性のとれた自然な位置合わせを行っている。

以上のように自然な見え方に関する研究はいくつか見られるものの、現実空間内で仮想物体を動かした際に現実物体が持っている素材の特性を考慮した仮想物体の自然な接触等についての研究はあまり見受けられない。

3. 提案手法

本研究では深層学習を利用した画像認識を用いることで素材の認識を行い、硬度や摩擦力の推定を行う。具体的には深度センサを内蔵したデバイスでカラー画像を取得し、これに深層学習を用いて物体の識別を行う。また、並行して認識した物体の位置に該当する部分を個々に抜き出したメッシュを作成する。次に識別された物体それぞれについて深層学習を用いて素材を判別し、作成したメッシュに素材を属性として付与することで、そのメッシュは金属、木製、布製などの情報を有することとなる。最後に生成されたメッシュすべてを1つにまとめ、現実空間の映像が透過するように重畳して表示することで、仮想物体がメッシュの影響を受けて跳ね方が変化したり接触した際の音が変わったりするなどの表現が可能となる。なお、硬度や摩擦は厳密に図るのではなく、万人が共通して感じることを前提として定義する。例えば布地であれば「柔らかい」「ざらつき」、金属であれば「固い」「つるつる」などである。事前にこのような定義をしておき、素材が付与されたメッシュに対して定義していた振る舞いを行うこととする。

4. システム構成

本研究では空間認識を行うための深度センサを内蔵したデバイスとして Kinect v2 を用いた。また、図.1 は本研究で構築したシステム全体の流れである。システムは大まかに物体認識、素材認識、認識した物体に対応したメッシュの生成、生成されたメッシュに素材の情報を付与、4つの処理から成る。本節ではこれら4つについて具体的な説明をしていく。

4.1 物体領域の抽出

初めに Kinect センサからカラー画像を取得する。この時取得する画像はのちに生成するメッシュとの位置合わせを考慮して、深度画像をもとにしたカラー画像を取得する。

次に物体の認識を行う。画像内に存在する物体を、深層学習を用いた画像認識によって抽出する。

本研究では実用的な学習済みモデルとニューラルネットワークのアルゴリズムを有するフレームワークである YOLO[3]を利用し、物体認識を行った。

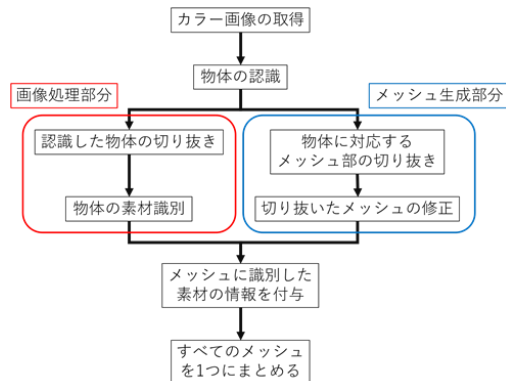


図.1 システムのフロー

図.2 は実際に処理を行った結果であり、認識した物体が何であるのか、その名称と領域が得られるので、結果を保存しておき、以降素材の認識、メッシュの生成の際に用いる。なお本研究は認識した物体がどのような素材で構成されているかを重要視しているため、正確な物体の把握までは行わない。故に画像内に物体がどのような位置にいくつ存在するのかを必要としているので、図.2 のようにテーブルが“Chair”と認識されているケースは問題ないとしている。

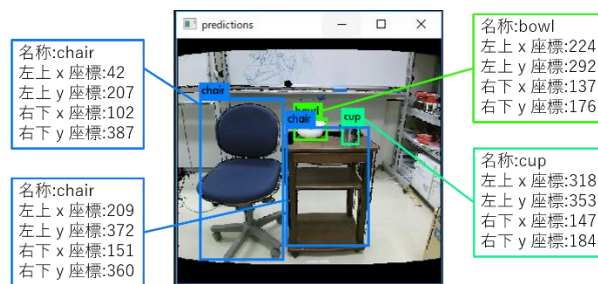


図.2 物体領域の抽出例

4.2 素材の推定

物体の認識の次に素材の推定を行う。まず、4.1 で取得した物体の位置を元に、カラー画像から物体領域を個別の画像として切り抜いていく。このとき、対象となる物体のみが映るように上下左右画素数の 10% をトリミングする。次にトリミングされた個々の画像に対し深層学習を用いて素材識別を行う。本研究では画像検索により素材の特徴が表れた画像を Web 上で収集し、深層学習用の学習データとして用いた。図.3 がその一例である。なお深層学習のフレームワークとしては Caffe[4]を用いた。

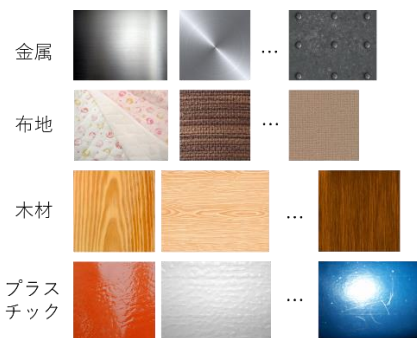


図.3 使用した学習用画像

4.3 点群の修正とメッシュ生成

4.2 の処理と並行してメッシュの生成を行う。認識された物体の位置を元に深度データから点群を取得する。図.4(a)のように取得したそれぞれの点群は一部重なる部分が存在する場合がある。メッシュにする際どちらかは不要になるため、この部分を削り取る必要がある。削り取る手法として kdtree[5]を使った点群探査で重なっている点を抽出し、残りの点群を修正後の点群として出力した。図.4(c)はその結果であり、綺麗に削除されているのが分かる。

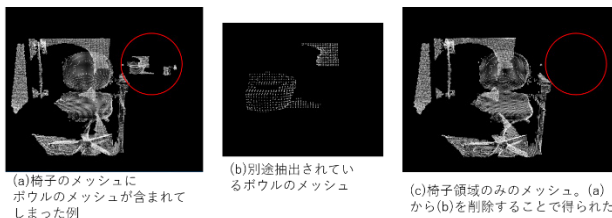


図.4 異なる点群で重なっている部分が存在する例

4.4 情報の付与

4.2 及び 4.3 において得られた素材の情報と個々のメッシュを対応付け、仮想物体と物理的な相互作用が可能な Mixed Reality のシーンを生成する。ここで生成するシーンは冒頭で述べたように単なる衝突判定だけではなく、現実世界に即した摩擦係数等の素材が持つ属性を付与したシーンである。

今回はメッシュに追加で与える属性として摩擦・反発係数および接触音を選択した。これらの属性はあらかじめ木材、金属、プラスチックといった代表的な素材のものをデータベースに登録し保存しておくようにした。4.2 の処理によって、抽出されたメッシュ群には既にそれらがどのような素材から構成されているかという名前が付け加えられている。前述したデータベースから該当する物体の係数を検索し、メッシュに与えてシーンとして生成する。

5. シーン生成処理の動作確認

前章までで述べたシーン生成処理の動作確認のため図.5 に対し一連の処理を施した。



図.5 シーン生成対象の環境

YOLO による物体領域の抽出結果は図.6 の通りである。椅子、机、ボウル、コップの4つの領域が検出されていることがわかる。ここで机が Chair となっているが、4.1 で説明した通り本研究では素材認識が重要で、物体名そのものの正確な認識はさほど必要ではない。

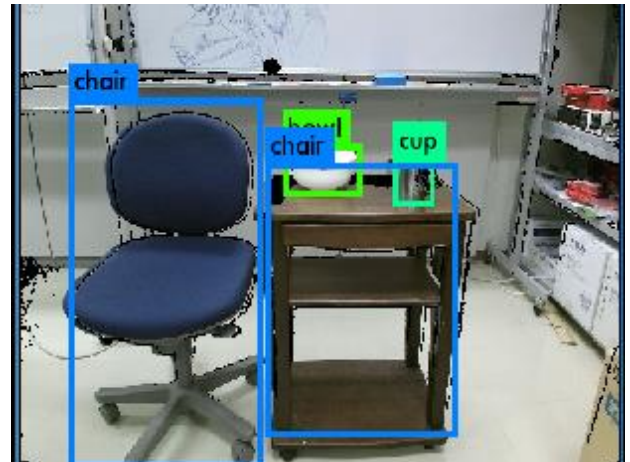
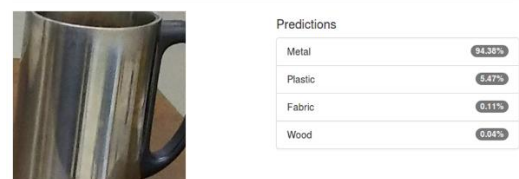


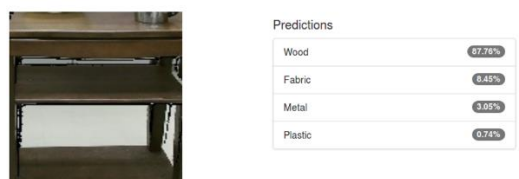
図.6 YOLO による物体領域の抽出結果

検出された物体領域それぞれに対し、4.2 で述べた素材認識を行った結果を図.7 に示す。コップ領域は「金属」、机領域は「木材」、ボウル領域は「プラスチック」として、正しく認識されたが、椅子領域は布製のクッションが大部分を占めているものの、「プラスチック」と認識されてしまった。

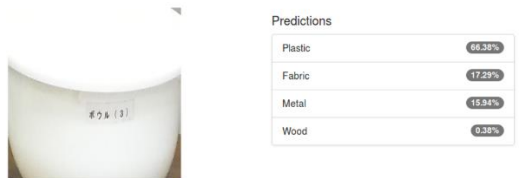
MCT_custom_100 Image Classification Model



MCT_custom_100 Image Classification Model



MCT_custom_100 Image Classification Model



MCT_custom_100 Image Classification Model

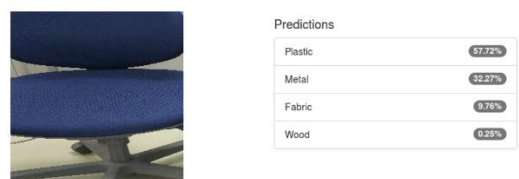


図.7 素材の認識結果

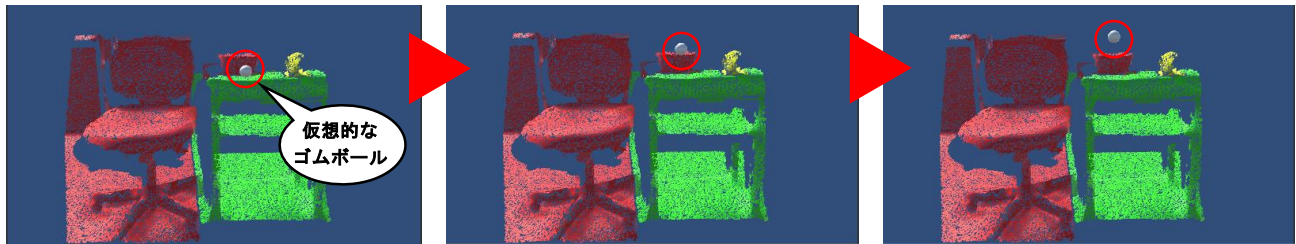


図.8 物理属性を付与したメッシュに、仮想物体のゴムボールを落下させた例

図.8 は生成された物理属性付きのメッシュに仮想的なゴムボールを落としたときの反発の様子を示している。赤いメッシュは「プラスチック」、緑のメッシュは「木材」、黄色のメッシュは「金属」の属性を付与されたメッシュである。木材でできた机から適切に上方に反発していることがわかる。また、紙面上では伝えられないが接触した際も設定した反発音が再生されており、既存の Mixed Reality のシーンよりもリアリティのあるシーンを生成することができた。

6. まとめと今後の展望

本研究では Mixed Reality 技術における仮想物体の表現を、従来の技術より現実感が深まるよう表現する手法を提案した。深度センサを内蔵したデバイスで実空間のカラー画像と点群を取得し、物体認識を行い現実空間内の物体の位置と数を把握するそして認識した物体に対し、深層学習を用いて素材の識別を行い、対応した点群へと反映、メッシュ化することで現実世界の物体の素材の情報を有したメッシュが完成する。このメッシュを透過し現実世界にオーバーレイすることで、仮想物体が接触可能となり、かつその振る舞いも素材情報の影響を受け変化する。これにより、現実物体と仮想物体のより現実的な表現が可能となった。

現時点では素材の推定のみで、その素材がどのような加工を受けているかまでは考慮されていない。例えば金属は表面が滑らかであると仮定はしたが、ものによってはざらざらした加工をしたものも存在する。同様に、ニスを塗った木製品は表面がなめらかに加工されている可能性がある。このように素材だけでなくその表面がどのように加工されているのかまで考慮することでさらに自然な表現が可能となってくると考えられる。

謝辞 本研究は JSPS 科研費 17K01160 の助成を受けたものです。

参考文献

- [1] 角田 哲也. Mixed Reality 技術を用いた文化財の復元 — 飛鳥京、江戸城、一乗谷 —. 遺跡整備・活用研究集会報告書, 平成 27 年度.
- [2] 稲葉 正樹, 池内 克史. 実世界環境モデル化による野外複合現実感のための頑健な位置合わせ. 東京大学

大学院 情報理工学系研究科 電子情報学専攻 修士論文, 平成 24 年度.

- [3] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi . You Only Look Once: Unified, Real-Time Object Detection. CVPR 2016, 9 May 2016.
- [4] Yangqing Jia, Evan Shelhamer, Jeff Donahue, Sergey Karayev, Jonathan Long, Ross Girshick, Sergio Guadarrama, Trevor Darrell. Caffe: Convolutional Architecture for Fast Feature Embedding. arXiv preprint arXiv:1408.5093. 2014.
- [5] P. Indyk, R. Motwani, P. Raghavan, and S. Vempala, Locality-preserving hashing in multidimensional spaces, In Proceedings of 29th ACM Symposium on Theory of Computing, pp.618-625, 1997.