

視点補間ネットワークによる実物体の質感再現 Reproduction of Material Appearance of Real Objects Using a Viewpoint Interpolation Network

西長 紗奈[†]
Sana Nishinaga

水野 瑛梨[†]
Eri Mizuno

星澤 知宙[‡]
Chihiro Hoshizawa

入山 太嗣[‡]
Taishi Iriyama

小室 孝[‡]
Takashi Komuro

小川 賀代[†]
Kayo Ogawa

1. はじめに

近年インターネット通販の普及に伴って、実際に商品を手にとらずに購入できるようになった。しかし多くの場合ユーザに提示されるのは数枚の写真だけであり、実際に商品を手にとって見るができないためその質感が分からない場合がある。

視覚的質感においては、視点位置による輝度の変化がリアリティに大きく影響する。実物体の質感再現においては、物体の光学特性を正確に取得する必要があるが、複雑な反射や透過をする人の肌や髪、果実の表面、ガラス等の物体における光学特性を画像から測定することは困難である。

この問題を解決するために、先行研究では、四端の4枚の画像から間の視点における補間画像を生成する視点補間ネットワークを提案し、金属物体の質感再現を行った[1]。この研究では、Blenderで作成した物体を一定の間隔で回転させた3DCG画像を用いて、訓練画像から入力画像まで3DCG画像のみを対象としている。

そこで、本研究では、実物体を対象としたときも視点補間ネットワークを用いて中間視点における補間画像が生成でき、質感再現が十分かどうか検討する。

2. 視点補間ネットワーク

本研究で使用する視点補間ネットワークはRonnebergerらが提案したU-Net[2]の原理を用いている。

U-Netは複数の入力画像を畳み込み、一枚の画像を生成するものであるが、本研究では両端の2枚の画像をそれぞれRGBに分けた6枚の画像と、位置情報を合わせた7枚を入力画像としている。また、入力と出力の画像サイズを揃えるために、各畳み込み層にゼロパディングを施している。視点補間ネットワークの概要を図1に示す。

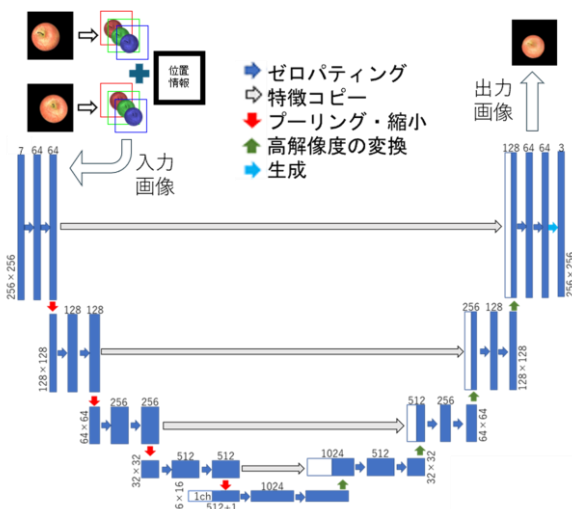


図1 視点補間ネットワーク

3. 実験

3.1 実物体の撮影

実物体の撮影環境を図2に示す。撮影環境は、広範囲から光が当たる室内で行った。背景は灰色または黒色の布を用いた。三脚にスライダーを取り付け、カメラは下向きに設置し、図2のピンク色の矢印の水平方向に微小移動させながら撮影を行った。撮影は、1cm間隔(微小角約2.5度)で1つの物体と設置角度の組み合わせに対し11枚撮影した。

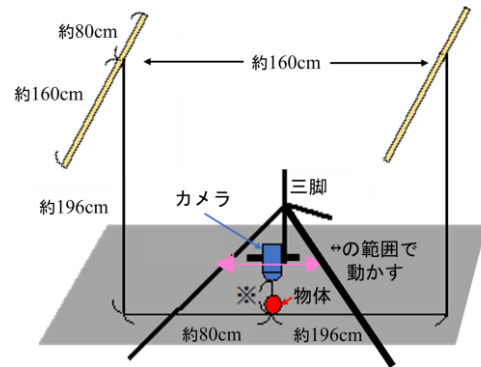


図2 撮影環境

撮影対象物体は、凹凸の有無、反射の有無など質感が異なるものを20種類(果物の食品サンプル、タワシ、電子時計など)用意した。また、プラスチック製の食品サンプルは食品の質感がよく再現されているものを使用した。

撮影した画像は、視点補間ネットワークによる対象物の誤認識の防止のため、物体が中央に配置されるよう1216px×1216pxの正方形にトリミングし、画像の背景を黒色にした。撮影した画像は20種類と100の設置角度の組み合わせで、合計2000枚である。

3.2 視点補間ネットワークの学習

視点補間ネットワークに学習させる訓練データは、Blenderで作成した3DCG画像を用いた。訓練画像の例を図3に示す。

3DモデルはThe Stanford 3D Scanning Repository[4]の(a) Stanford Bunny、(b) Happy Buddha、(c) Dragon、(d) Lucy、(e) Asian Dragon、(f) Thai Statueの6種類である。材質は陶器や石鹸、人肌に見られる表面下散乱材質を使用した。色やサイズはランダムに変更している。撮影画像と同様のカメラと蛍光灯の位置を設定し、1cmずつ移動させた11枚の画像を作成した。入力画像として両端の2枚の画像、正解画像として両端の間の9枚の画像を3Dモデル1種につき、1000組用意した。よって11枚と6種類と1000組を組み合わせ合わせた合計66000枚作成し、使用した。

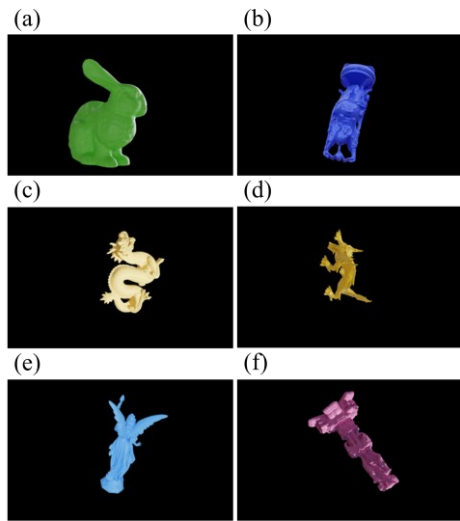


図 3 訓練画像に用いた 3D モデル

4. 生成結果

撮影した実物体を入力画像とし、視点補間ネットワークを用いて補間画像の生成を行った。生成された食品サンプルのライムの補間画像の結果を図 4 に、食品サンプルのスターフルーツの補間画像の結果を図 5 に示す。

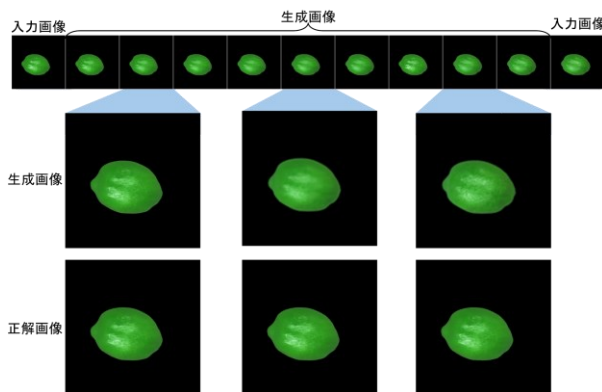


図 4 ライムの補間画像の結果

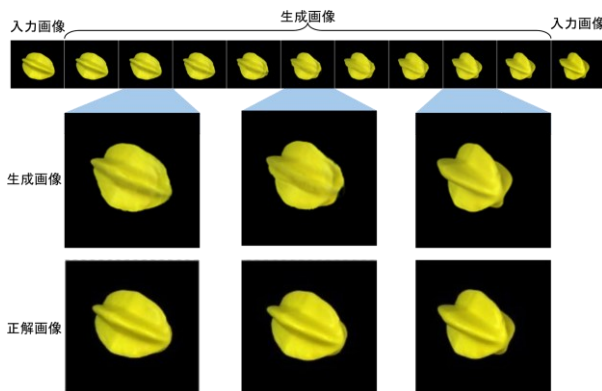


図 5 スターフルーツの補間画像の結果

両端の 2 枚の画像が撮影した画像かつ入力画像であり、両端の除いた 9 枚の画像が生成画像である。ライムの生成画像より、光源による反射の移動が確認できる。また、スターフルーツの生成画像より、視点位置による物体の形状が変化している物体が生成されていることが確認できる。

また、質感の違いがわかりやすいタワシと電子時計の補間画像の例を図 6 に示す。

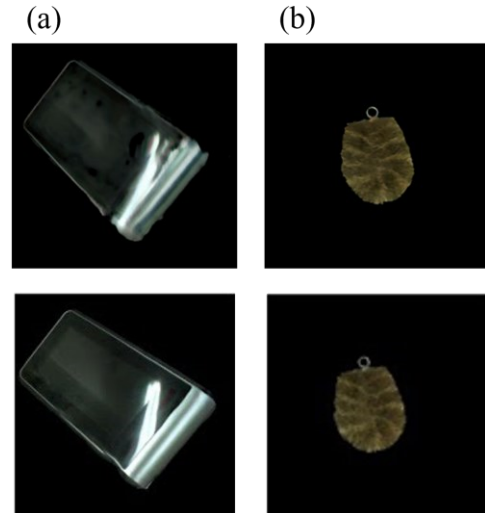


図 6 質感がわかる補間画像の例

図 6 の上段は最も視点移動が大きい位置である生成画像の中央の画像であり、対して下段は端の画像である。(a)の電子時計は表面が反射する物体であり、補間画像でも蛍光灯の反射が再現されている。これに対し、(b)のタワシは反射しておらず、加えて棘の細かな質感など、反射の違いや質感が再現されている。学習に用いた CG 画像は表面下散乱材質のみであるにも関わらず、質感の違いを再現できていることは、この視点補間ネットワークの精度が高いことを示している。

5. まとめと今後の課題

撮影した実物体を入力画像とした補間画像の生成を行い、視点位置による物体の形状や反射の変化を反映した画像が生成され、実物体の質感が再現されたことを確認した。今後は、生成画像に対する客観的評価を行う。

参考文献

- [1] 星澤知宙, 小室孝, “質感再現のための視点補間ネットワーク”, 第 28 回画像センシングシンポジウム (SSII 2022), 講演論文集, IS3-09, SO3-09 (2022)
- [2] Olaf Ronneberger, Philipp Fischer, Thomas Brox, “U-net: Convolutional networks for biomedical image segmentation”, in International Conference on Medical image computing and computer-assisted intervention, Springer, p234-241 (2015)
- [3] Edmund Optics, “EO-23122M モノクロ USB 3.0 カメラ”, <https://www.edmundoptics.jp/p/eo-23122m-monochrome-usb-30-camera/3209/>, (Accessed: June 9, 2024)
- [4] Stanford University Computer Graphics Laboratory, "The Stanford 3D Scanning Repository", <http://graphics.stanford.edu/data/3Dscanrep/>, (Accessed: June 9, 2024)

† 日本女子大学大学院理学研究科 Graduate School of Science, Japan Women's University

‡ 埼玉大学大学院理工学研究科 Graduate School of Science and Engineering, Japan Saitama University