

3 システムの概要

提案システムはカスタマーとワーカーがともにブラウザ上で操作を完結させるウェブシステムとして構築されている。カスタマーはイラストや写真といった画像をアップロードしたのち、矢印によって画像中の物体の前・横・上方向を指示し、また生成されるモデルを構成すべきパーツ数の下限を入力する。システムはクラウドソーシングプラットフォーム (CrowdFlower¹) にジョブオファーを作成し、スケッチを行うワーカーを60名 (20名×3方向) 募集する。必要枚数のスケッチが集まると、それらを用いてシステムは3Dモデルを自動生成し、カスタマーに提示する。カスタマーがモデルの品質に満足した場合は処理が完了し、不満足であった場合は品質向上のためにシステムがさらに多くのスケッチを集める反復的なワークフローを開始する。

3.1 スケッチタスク

ワーカーは図1に示すようなページに誘導される。左側に入力画像および投影方向を示す矢印が提示されており、右側のドローイングインタフェースを用いてマウス操作により基本図形 (四角形・楕円形) を変形・配置して投影図をスケッチとして作成する。ワーカーには、元の画像では他のパーツに遮蔽されて部分的にまたは完全に隠れているものも含めて、物体に含まれる“すべての”パーツを記述するように指示が与えられ、カスタマーに指定された最小パーツ数の基本図形を配置するとスケッチの提出が許される。

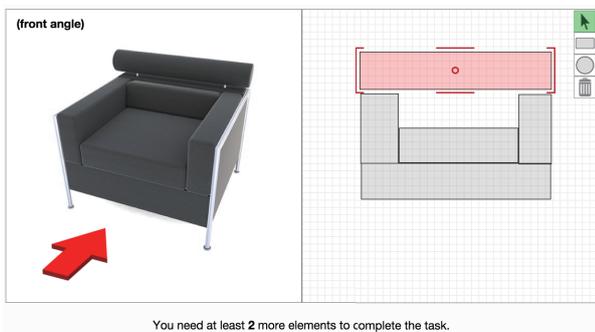


図 1: スケッチインタフェース

すべてのスケッチの提出を完了したワーカーには36セントの報酬が支払われ、後述のピアレビューによって提出スケッチの品質が上位20%に評価されたワーカーには、追加報酬としてさらに18セントが与えられる。評価基準は「正しい方向から描かれていること」「対象物体に含まれるすべてのパーツが描かれていること」「要素の配置が正確であること」「対象物体に含まれない要素が描かれていないこと」であると指示されている。報酬金額は、アメリカ

合衆国における最低賃金 (\$7.25/hr) および想定作業時間 (3分間) に基づき設定を行った。

3.2 ピアレビュータスク

ワーカーによって提出されたスケッチにはでたらめなものやタスク指示を誤解したものが含まれるため、品質管理のため、スケッチワーカーと同様の方法で雇用されたワーカーによって評価が与えられる (図2)。レビューワーカーにはスケッチタスクの説明が与えられ、20名のワーカーから提出されたスケッチの品質を、上述の評価基準によって7段階 (1:最低, 7:最高) で評価することが求められる。

各スケッチの最終的な評点は20名のレビューアークから与えられたスコアの平均値により決定される。すべてのレビューワーカーには一律に24セントの報酬が与えられる。

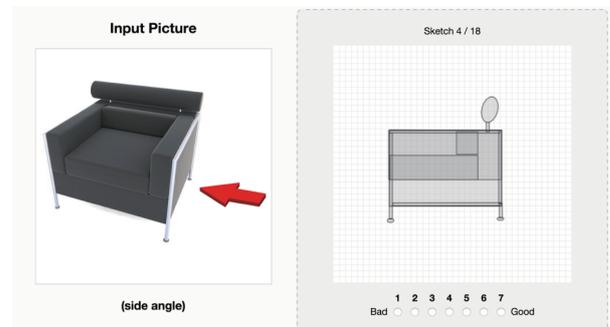


図 2: レビューインタフェース

3.3 反復的ワークフロー

スケッチタスクのみを用いた予備実験により、およそ40%のスケッチしか正しい方法で描かれなかったことや、正しく描かれたスケッチであっても完成度が低く、細かなパーツなどは無視されやすいため立体生成に用いるには品質が不十分であるという問題が判明した。前者については、無効なスケッチには不正行為によるでたらめなものも含まれる一方で、半数近くが注意深く描かれているものの描画方向が誤っているなどタスクの誤解に基づくものであることも確認された。これらのワーカーに対しては、適切な見本を示すことで誤解を防ぐことが期待できる。後者については、最低限のタスク要請を満たす以上の努力量を費やすインセンティブが存在しないことが完成度の低さの要因となっていると考えられるので、提出物の品質に応じたボーナスを示唆することで改善が期待できる。

こうした観察から、一定数のワーカーからスケッチを集め、品質の高いものに追加の報酬を与えることで競争を促しつつ、これらを次に集めたワーカーへの手本として示すことでタスク理解も促進させ (暗黙的な協力)、両問題を

¹<https://www.crowdflower.com/>

同時に解決することとした。具体的には次のような反復的なワークフローとして実現される (図 3)。

1. 見本を示さずに 20 枚のスケッチをワーカーから収集する。
2. ピアレビューにより各スケッチに評点を付ける。
3. 上位 20% (4 枚) に選ばれたスケッチの提出者に 18 セントの報酬を与える。
4. 三次元モデルを生成し、カスタマーに提示する。品質が十分と判断された場合は処理を終了する。
5. さらに 20 枚のスケッチを新たなワーカーから収集する。この際、スケッチインタフェース上に上位スケッチを手本として提示する。
6. 2 に戻る。

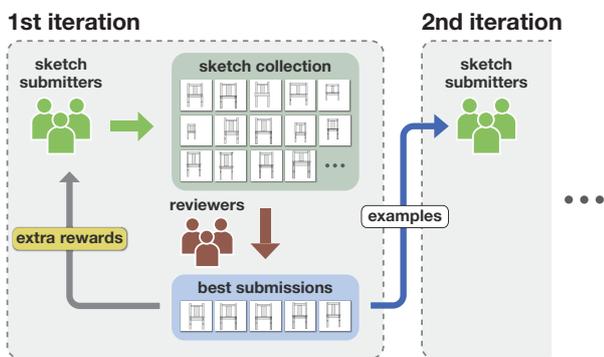


図 3: 反復的ワークフロー

これは、遺伝操作 (以前の提出物を参考としたスケッチ作成) および選択 (ピアレビュー) の双方を人間 (ワーカー) が行う一種の遺伝的アルゴリズム [15] として解釈できる。各段階のスケッチ及びレビュータスクのワーカー募集は独立に行われるので、原則的に異なるワーカーが各タスクには参加することになるが、現在は一人のワーカーが複数回のタスクに参加することを禁止してはいない。

4 立体生成アルゴリズム

4.1 有効スケッチの選出

反復的ワークフローによる品質改善を経ても提出スケッチには誤ったものが一定数含まれているため、正しく描かれたスケッチのみを立体復元に用いる必要がある。ここでは、正しいスケッチは必然的に互いに見た目が似通っており、誤ったスケッチはそれぞれ大きく異なる見た目をしているという観察から、スケッチの類似性に基づくクラスタリングによる選出手法を提案する。

まず、ある視線方向について収集したスケッチのペアが互いにどれだけ似通っていないかを示す、距離行列を計算する。ここでは、平行移動や詳細部の描き込みに左右されないことを考慮し、各スケッチのバウンディングボックスを $128\text{px} \times 128\text{px}$ の正方形に揃えるようスケールした上で、修正ハウスドルフ距離 (MHD) [6] を計算することで距離指標として採用した。

次いで、距離行列に基づいてスケッチを分類する。ここではクラスタの個数が事前に未知であることや、距離行列のみが与えられており平均値などが定義できない状況から、Medoidshifts 法 [20] をクラスタリング手法として採用した。クラスタリング結果のうち、最大スケッチ数を含むクラスタを、正しいスケッチ (有効スケッチ) のみから構成されるものとして採用する。実際に、椅子の入力画像に関する横方向のスケッチ群に対して提案手法で距離計算およびクラスタリングを行った結果を図 4 に示す。

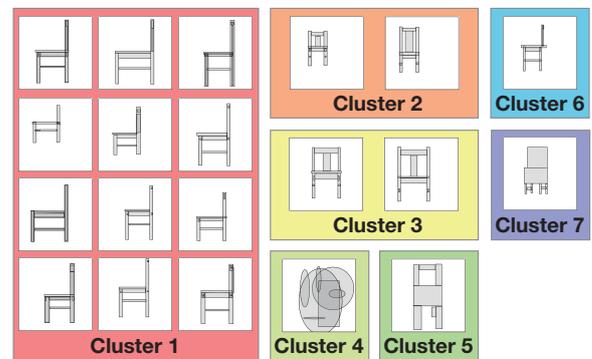


図 4: 椅子画像の横方向スケッチ群に対するクラスタリング結果の例。Cluster 1 には適切に描かれたスケッチが全て集約されており、また誤ったスケッチは含まれていない。

4.2 三面図の生成

各視線方向について、有効スケッチを統合してクリンナップされた三面図を生成する。まず、すべての有効スケッチから各基本図形 (長方形・楕円形) のみを抽出した画像を生成し、これらの中で上述の手法で距離行列を計算してクラスタリングを行う。この際、同じ種類の基本図形 (長方形または楕円形) に関してそれぞれクラスタリングする。これにより、各スケッチに含まれる図形要素を、互いに同じパーツに対応するものごとのクラスタにまとめることができる。含まれる要素数が閾値 (4 に設定) を下回るクラスタはノイズとみなして無視し、残ったクラスタごとにサイズ・回転・位置の平均値を計算してパーツ形状のコンセンサスを求める。各パーツのコンセンサスを集めることで、クリンナップされた投影図が生成される (図 5)。これを三視線方向についてそれぞれ行うことで、三面図を得ることができる。

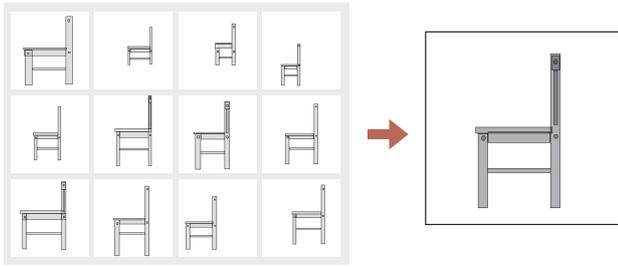


図 5: 12 枚のスケッチの統合例

4.3 3D モデル生成

最後に、三面図に含まれる基本図形同士の対応関係を推測し、単一のパーツに対応すると考えられる各三つ組から三次元プリミティブを生成することで3Dモデルを得る。まず、三図面内の各基本図形が縦横に占める区間を計算する。二つの図面に含まれる基本図形について、図面の接する方向（前方向と横方向の図面なら、垂直方向）の占有区間端部間の誤差二乗和をコストとして定義する。三つの図面から得られるすべての三つ組について、含まれるペアのコストの和が閾値（キャンバスサイズ 400px において、1000px² を経験的に採用）を下回るものを、単一のパーツを三方向から見た図形の組であると見なし、対応する三次元プリミティブを生成する（図 6）。現在はスケッチの基本図形は長方形と楕円のみに対応しており、生成できる三次元プリミティブは直方体、楕円柱および楕円体に限られる（図 7）。

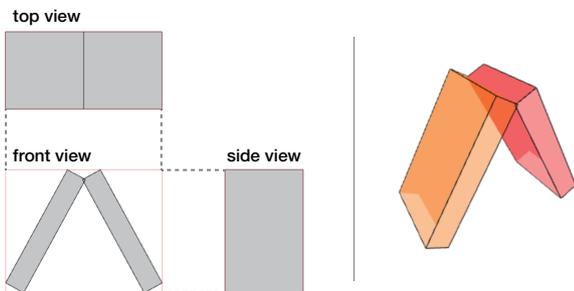


図 6: 三面図からの 3D モデル生成例

| view 1 | view 2 | view 3 | 3D primitive |
|-----------|-----------|-----------|--------------|
| rectangle | rectangle | rectangle | cuboid |
| rectangle | rectangle | ellipse | cylinder |
| ellipse | ellipse | ellipse | ellipsoid |

図 7: 現在対応している三次元プリミティブの一覧

5 結果

本システムを用いて様々な入力画像（図 8）について実際に 3D モデル生成を行った結果を図 9 に示す。出力モデルを見ると、対象物体に含まれる主要なパーツは概ね正しい配置で生成されており、ソファのレッグカバーなど細かな要素も一部再現されていることが確認できる。また、椅子の側板やソファの後ろ脚など、入力画像には陽に含まれない要素も、ワーカーが実世界知識から存在を推測して描画したことで、合理的な形で生成されている（図 10）。



図 8: 入力画像例

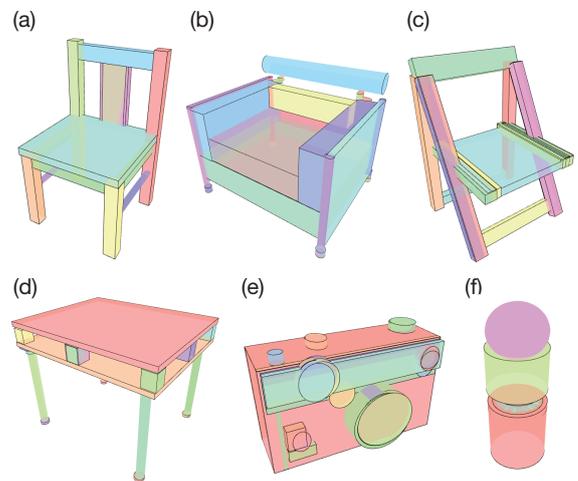


図 9: 3D モデルの生成例

各画像の生成に掛かった時間および金銭的成本を表 1 に示す。入力画像 (b) に関して Lancers² を通じて雇用したプロフェッショナルに依頼した際（結果モデル：図 11）にはタスク完了までに約 24 時間の所要時間と \$45 相当の金銭が必要とされたことと比較すると、提案手法は金銭面的コストおよび生成モデルの細部のクオリティではプロへの依

²<http://www.lancers.jp/>

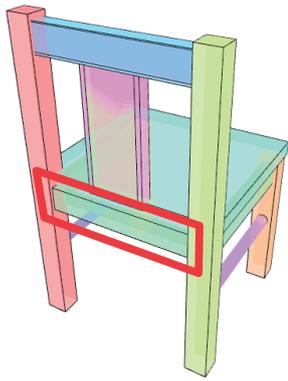


図 10: 生成された椅子 3D モデル (図 9a) の背面. 入力画像には陽に含まれていない側板が確認できる.

頼に劣るものの, 所要時間ではメリットがあることが分かる. また, 提案手法は作業依頼のための対人コミュニケーションが不要であるため, カスタマーの労力は著しく小さいと言える.

表 1: 各入力に対する時間・金銭コストの一覧

| 入力 | 時間 | 所要金額 | 反復回数 | スケッチ数 |
|----|--------|----------|------|-------|
| a | 3h 34m | \$228.9 | 5 | 300 |
| b | 3h 03m | \$228.9 | 5 | 300 |
| c | 1h 52m | \$137.34 | 3 | 180 |
| d | 1h 10m | \$91.56 | 2 | 120 |
| e | 1h 27m | \$91.56 | 2 | 120 |
| f | 0h 45m | \$45.78 | 1 | 60 |



図 11: プロフェッショナルによるモデリング結果.

提案の反復的ワークフローの効果を示す例として, 入力画像 (a) について, 1・3・5 回目の反復での収集スケッチ (20 枚 × 3 方向 = 60 枚) のみに基づく 3D モデル生成結果の比較を図 12 に示す. 初期のスケッチからはごく一部のパーツしか生成できていないものが, 回を重ねるごとに多くのパーツを正しく生成できるように品質が向上していることが確認できる.

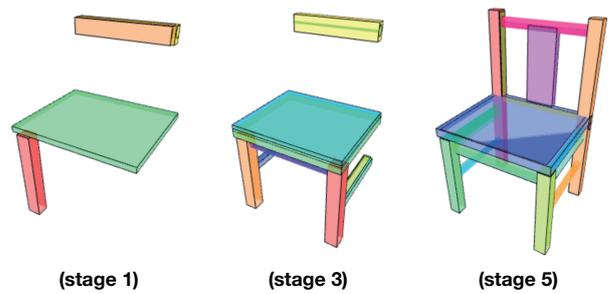


図 12: スケッチ品質の反復的改善による生成結果の比較

図 13 に椅子およびソファ入力画像に関して, 各方向の反復ごとの有効スケッチ割合の変化を示す. 2 回目の反復ではワーカーにはじめて手本が示されるが, この際に全ての入力画像について有効スケッチ割合が増加していることが確認できる. その後は入力によって増減があるものの, いずれも 5 回目の反復では約 70–90% の値が得られており, ワークフローの効果が確認できる.

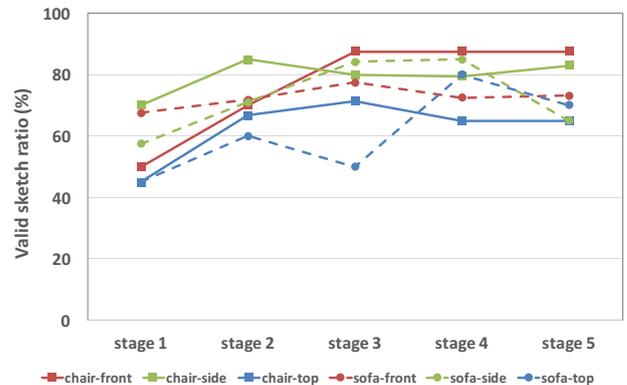


図 13: 有効スケッチ割合の反復的改善

また本手法は, ワーカーが一意的に描かれている物体の立体配置を解釈できる限りは, 写真のほかにイラストレーションや手書きのスケッチなどを入力に利用してもアルゴリズムやシステムの変更なく三次元モデルを生成することができる. 図 14 に手描きのタンスのスケッチおよび, これを入力として三次元モデルを生成した結果を示した. このモデルの生成は一回のスケッチ収集のみ, およそ 15 分間の所要時間で完了したが, 全ての描かれたパーツが正しく配置されていることが確認できる.

6 議論

6.1 想定ユーザー

提案手法の想定される利用例としては, 専門的な三次元モデリング技能を持たない一般ユーザーが, 3D プリンティ

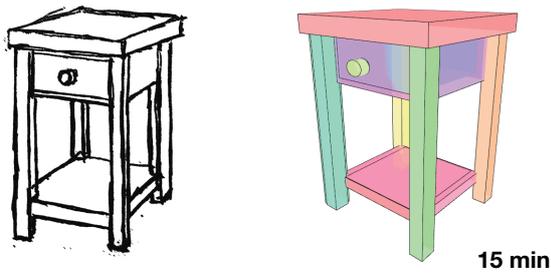


図 14: タンスの手描きスケッチ (左) および生成結果 (右) .

ングなどを利用してオリジナルのオブジェクト (例: 歯ブラシ立てや電話スタンドといった日用品, ドールハウスの家具, ゲームトークンなどの玩具) を作成する際, 簡便なスケッチのみを作成するだけでモデリングを可能にするという応用が考えられる. 安価でかつ品質の高い 3D モデリングサービスを提供しているプロフェッショナルワーカーには数に限りがあるが, 本技術では事実上無尽蔵のマイクロワーカーの資源が活用できるため, 需要が増大しても常に一定品質のサービスを提供することができる.

この他, インディーズゲームのリソース作成など, ハイエンドの品質が要求されない 3D モデルを多数作成したい場合にも同様にマイクロワーカーの資源を活用する余地があると考えられる. また, オンラインゲーム等に本システムを組み込むことで, ユーザーオリジナルのアイテム作成を可能にすると同時に, タスクの対価をゲーム内貨幣によって支払うことで, 実質的な金銭コストも引き下げられる可能性がある.

6.2 簡易モデリングソフトとの比較

SketchUp³をはじめとしたインタフェースが簡易化された 3D モデリングツールを用いれば, 専門技能を持たないユーザーでも 3D モデルを作成することも可能である. ただし, そうしたソフトウェアでも 3D インタフェース内のオブジェクト移動や回転, モード切替などの専用操作は必要であり, 使用方法を覚えるためには一時間から数時間程度の学習時間が要求される. この点, 提案手法はカスタマーとワーカーのどちらも 3D 操作を一切行う必要なく, 慣れ親しんだ伝統的な 2D マウス操作のみで作業が完結するという相違がある. 本研究では, こうした学習時間および操作時間の低減により, 数分間しかワーカーのエンゲージメントが得られないマイクロタスク型クラウドソーシングを, 3D モデリングという複雑かつ創造的な課題に対しても活用できる可能性を示した.

6.3 制約点と研究課題

本研究は, 3D モデリングの簡易タスクへの分割およびクラウドソーシングワークフローの提案によるモデリングパイプラインの構築に主眼を置いたため, ユーザー (カスタマー・ワーカー) 視点におけるシステム評価およびインタフェースの改善についてはさらなる研究が必要である. 例えば, 現在は不正なスケッチ投稿を排除するためにカスタマーにパーツ数の見積もりを入力することを求めているが, これは “1-5,” “6-10” のように, より大雑把な指定で十分である可能性がある.

また, 現在のシステムは三種類の 3D プリミティブの生成および XYZ いずれか 1 軸に関する回転のみに対応しており, 任意の軸回りの回転や, 押出・膨張, CSG 操作や自由曲線といった高度な形状要素の生成はできない. 既存のスケッチベースドモデリング等の手法を本システムにいかに関組み込み, 多様な形状生成を可能にできるかは今後の課題である. この際, 例えばエッジの丸め処理の度合いや要素の均等配置等についてワーカーに判断を委ねるなど新たなクラウドタスクを追加することで, 基本形状の生成に留まらず 3D モデリングの様々な側面にマイクロタスク型クラウドソーシングを活用することができるだろう.

7 結論

本研究では, マイクロタスク型クラウドソーシングを用いた 3D モデリングシステムを提案し, 多数のワーカーからスケッチを収集することで実際に様々な画像から 3D モデルを生成できることを示した. 今後の発展として, 対応 3D プリミティブの追加や要素の均等配置など, より高度なモデリングプロセスへの本手法の拡張が考えられる.

参考文献

- [1] L. von Ahn, “Human Computation,” PhD thesis, Carnegie Mellon University, 2005.
- [2] L. von Ahn, and L. Dabbish, “Labeling Images with a Computer Game,” Proc. CHI 2004, pp. 319–326, Vienna, Austria, Apr.2004
- [3] M. S. Bernstein, G. Little, R. C. Miller, B. Hartmann, M.S. Ackerman, D.R. Karger, D.Crowell, and K. Panovich, “Soylent: A word processor with a crowd inside,” Proc. UIST 2010, pp. 313–322, New York, USA, Oct.2010
- [4] T. Chen, Z. Zhu, A. Shamir, S.-M. Hu, and D. Cohen-Or, “3-sweep: Extracting editable objects

³<https://www.sketchup.com/>

- from a single photo,” *ACM Trans. Graph.*, vol.32, no.6, pp. 195:1–195:10, Nov.2013.
- [5] S. Cooper, F. Khatib, A. Treuille, J. Barbero, J. Lee, M. Beenen, A. Leaver-Fay, D. Baker, Z. Popovic, and Foldit players, “Predicting protein structures with a multiplayer online game,” *Nature*, vol.466, no.7307, pp. 756–760, Aug.2010.
- [6] M.-P. Dubuisson and A. Jain, “A modified hausdorff distance for object matching,” *Proc ICPR 1994*, pp. 566–568, Jerusalem, Israel, 1994.
- [7] D. Eigen, C. Puhrsch, and R. Fergus, “Depth Map Prediction from a Single Image using a Multi-Scale Deep Network,” *Proc. NIPS 27*, pp. 2366–2374, 2014.
- [8] U. Gadiraju, R. Kawase, S. Dietze, and G. Demartini, “Understanding malicious behavior in crowdsourcing platforms: The case online surveys,” *Proc. CHI 2015*, pp. 1631–1640, Seoul, Korea, 2015.
- [9] Y. Gingold, A. Shamir, D. Cohen-Or, “Micro Perceptual Human Computation for Visual Tasks,” *ACM Trans. Graph.*, vol.31, no.5, pp.119:1–119:12, Sep.2012.
- [10] Y. Gingold, E. Vouga, E. Grinspun, and H. Hirsh, “Diamonds from the rough: Improving drawing, painting, and singing via crowdsourcing,” *Proc. HCOMP 2012*, Toronto, Canada, Jul.2010.
- [11] E. Guillou, D. Meneveaux, E. Maisel, K. Bouatouch, “Using vanishing points for camera calibration and coarse 3D reconstruction from a single image,” *The Visual Computer*, vol.16, no.7, pp. 396–410, 2000.
- [12] Q. Huang, H. Wang, and V. Koltun, “Single-view Reconstruction via Joint Analysis of Image and Shape Collections,” *ACM Trans. Graph.*, vol.34, no.4, pp.7:1–87:10, Aug.2015.
- [13] T. Igarashi, S. Matsuoka, and H. Tanaka, “Teddy: A Sketching Interface for 3D Freeform Design,” *Proc. SIGGRAPH 1999*, pp. 409–416, Los Angeles, USA, Aug.1999.
- [14] A. Kittur, B. Smus, S. Khamkar, and R. E. Kraut, “Crowdforge: Crowdsourcing complex work,” *Proc. UIST 2011*, pp. 43–52, Santa Barbara, USA, 2011.
- [15] A. Kosorukoff, “Human based genetic algorithm,” *Proc. SMC*, pp. 3464–3469, Tucson, USA, Oct.2001.
- [16] Y. Koyama, D. Sakamoto, and T. Igarashi, “Crowd-powered parameter analysis for visual design exploration,” *Proc. UIST 2014*, pp. 65–74, Honolulu, USA, Oct.2014.
- [17] B. Mason and S. Thomas, “A million penguins research report,” *Institute of Creative Technologies, De Montfort University, Leicester, UK*, 2008.
- [18] C. H. Park, K. Son, J. H. Lee, and S.-H. Bae, “Crowd vs. crowd: Large-scale cooperative design through open team competition,” *Proc. CSCW 2013*, pp. 1275–1284, San Antonio, USA, 2013.
- [19] A. Rivers, F. Durand, and T. Igarashi, “3d modeling with silhouettes,” *ACM Trans. Graph.*, vol.29, no.3, pp. 109:1–109:8, July 2010.
- [20] Y. Sheikh, E. Khan, and T. Kanade, “Mode-seeking by medoidshifts,” *Proc. ICCV 2007*, pp. 1–8, Rio de Janeiro, Brazil, 2007.
- [21] L. Yu, and J.V.Nickerson, “Cooks or Cobblers?: Crowd Creativity Through Combination,” *Proc. CHI 2011*, pp. 1393–1402, Vancouver, Canada, 2011.
- [22] R. C. Zeleznik, K. P. Herndon, and J. F. Hughes, “SKETCH: An Interface for Sketching 3D Scenes,” *Proc. SIGGRAPH 1996*, pp. 163–170, New Orleans, USA, 1996.