

運動視差を用いた直感的なマルチタッチインタラクション

原 田 一 馬[†] 菅 野 裕 介[†] 佐 藤 洋 一[†]

本研究では、運動視差を用いた自由な視点の変更とマルチタッチデバイスを用いた直感的な物体操作を両立した新しいインタラクションを提案する。マルチタッチによる対象の3次元操作を実現しようとした場合、2次元平面での入力をもとに3次元空間内の対象物体を操作しなければならないため、何らかの制約を設けなければ曖昧性が残ってしまう。本研究では頭部情報および運動視差を用いてこの曖昧性を排除することでマルチタッチによる3次元空間の操作を可能にし、既存研究との評価実験により有効性を示す。

Direct Multitouch Interaction in Motion Parallax

KAZUMA HARADA,[†] YUSUKE SUGANO[†] and YOICHI SATO[†]

We introduce a new interaction framework that utilizes both multi-touch inputs and motion parallax created with free head motion. When we consider 3D manipulation using multi-touch inputs, it becomes necessary to incorporate some constraints to cope with the discrepancy between 2D inputs on multi-touch devices and 3D control of a target object. In this paper, we propose the operation of multitouch three-dimensional space by eliminating this ambiguity using head information and motion parallax, and show the effectiveness of evaluation experiments.

1. 序 論

近年マルチタッチ可能なディスプレイが広く普及し、指でディスプレイを直接接触するという直感的な操作を可能にした。このようなデバイスを用いて2次元操作を行う場合、入力と操作対象の空間が2次元で一致しているため、面内平行移動、拡大縮小などの操作を直感的に行うことが可能である。これに対して、マルチタッチによる3次元操作を考えた場合、入力空間が2次元平面であるため、入力と操作をどのように結び付けるかが必ずしも自明ではない。ジェスチャーの割り当てによりこの問題を解決する研究が数多く行われているが、ジェスチャーの一般性や操作精度にはまだ課題が残っている。

一方、物体の3次元操作をする場合、その空間をいかに人間にとって認知しやすい形で表示するかという問題も重要な要素である。人間が立体的に知覚する仕組みは、主に両眼視差と運動視差の2つに分けられるが、特に運動視差は立体知覚をする上で空間把握に大きく寄与している。運動視差による3次元表示を実

現するために必要な情報はユーザの視点位置のみであるため、カメラを用いれば容易に実現できる。マルチタッチデバイスと運動視差による3次元表示の組み合わせで直感的な3次元操作と空間把握を両立をすることが出来れば、デバイスの入手性やデバイスを非装着性などの大きな優位性がある。

運動視差が伴う環境下でマルチタッチインタラクションを実現するためには、解決すべき2つの重要な課題がある。

- (1) 操作対象物体が指に追従しないこと (図1)
- (2) 意図しない方向へ回転運動すること (図2)

提案システムは、第一の問題にはタッチ点の深さにおける視差を最小化する方法で、第二の問題には頭部姿勢の情報を用いて回転方向を決定する方法で解決を試みる。

以下本論文では、初めに関連した研究を整理し、3次元物体の操作における課題を検討する。次にマルチタッチインタラクションと運動視差の2つの技術的要素を説明し、本研究が提案する手法の詳細を提示する。さらに評価実験にて本研究の有用性を示した上で、最後に今後の課題と将来的な展望を述べる。

[†] 東京大学生産技術研究所
Institute of Industrial Science, the University of Tokyo

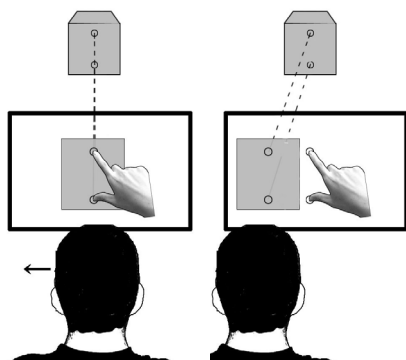


図 1 タッチ点のずれ、視点を変更した際、物体がディスプレイに投影される位置が変化する。

Fig. 1 Gap of touched positions. Projected position of the target object can change according to head motion.

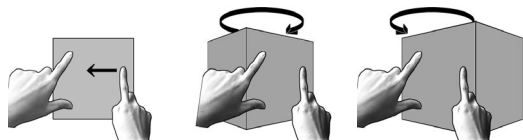


図 2 あいまいな回転。回転方向が解決できないため、意図しない回転が生じる。

Fig. 2 Ambiguous rotation. Target object can rotate to unintended directions.

2. 関連研究

マルチタッチデバイスにおいては入力空間が 2 次元平面であるため、奥行き方向の座標をいかに指定するかが重要である。マルチタッチデバイスを用いたインタラクションの一つとして Hancock らは、3 次元空間において 3 本までの指を用いて 3 次元物体を操作する方法¹⁾を提案した。しかしながら彼らの方法は 2 次元平面では認識しにくい奥行き情報を捨てることにより実現しており、一般的な 3 次元空間での操作には適応できない。Martinet らの研究²⁾では、XYZ 各軸からの平行投影の 3 面図が用いられている。3 面図により 3 次元空間を把握し、Z 軸方向の座標を直接指定することができるが、1 画面のみで 3 次元的操作を行うことは困難である。

タッチ平面のみで奥行き方向を取得する研究として、Reisman らは 2 次元タッチ情報から物体上のタッチ点がどのように変化するか推定する手法を提案した³⁾。この手法では特定のジェスチャーに頼らない移動や回転が可能である一方、回転運動に関して曖昧性が生じている。この Reisman らの研究で存在する回転の曖昧性を排除するために、Martinet らは移動と回転を分離する研究⁴⁾を行った。しかしながら、移動と回転が同時にできなくなるという別の問題が生じている。

このようにタッチ平面のみで 3 次元空間を扱う場合、奥行き方向に対して曖昧性が生じるため、指のみを用いて任意の 3 次元操作を実現することは困難である。

一方、平面のタッチディスプレイに頼らない方法で、奥行き情報を取得する研究も行われており、タッチデバイスを球体状にした研究⁶⁾や、立方体状にした研究⁷⁾がある。また、タッチディスプレイに他のデバイスを組み合わせる研究としては、トラックボールを用いた例⁵⁾、タッチディスプレイに圧力センサーを仕込んだ例⁸⁾、モバイルディスプレイを用いた例¹⁰⁾などが挙げられる。これらは別のデバイスにより奥行き情報を取得することでタッチ平面の曖昧性が生じる問題を回避しているが、使用するデバイスにより操作できる空間や物体が限られてしまう。

Valkov らの研究⁹⁾では、両眼視差が発生する状況におけるタッチの影響が検証されている。また Valkov らは、両眼視差と WIM と呼ばれる VR システムを組み合わせさせた研究¹¹⁾も行い、視差を用いコンテンツを立体的に表示した場合、空間の奥行き方向を認知しやすいという結論を得ている。その他に、マルチタッチと視差を用いた立体的な視覚化を組み合わせさせた研究として、Rivière らの研究¹²⁾が挙げられる。しかしこの研究は平面的な操作のみ (XY 移動, Z 回転) を対象としており、全ての軸に対する 3 次元操作は出来ない。Hechet らは鏡ベースのディスプレイシステムによる視差表現を実現したシステム上で、立体的な物体の移動および回転をする研究¹⁴⁾を行った。この研究では物体を直接タッチせず、マニピュレータを用いて物体を操作している。これらは視差を立体表示にのみ使用しており、マルチタッチによる 3 次元操作には生かされていない。以上のように、タッチ平面の奥行き方向の曖昧性を解消するために視差を用いることの有効性が従来研究により示唆されている一方で、こうした環境下でのマルチタッチインタラクションの設計に関しては十分な議論がなされていない。

3. 提案システム

本章では、運動視差を伴う環境下の直感的なマルチタッチインタラクションを実現するために必要な技術について述べる。最初に運動視差による 3 次元表示の実現手法と Reisman らの手法³⁾を元にした 3 次元操作アルゴリズムについて整理した上で、より直感的な操作を実現するための提案手法について述べる。

3.1 運動視差環境下におけるマルチタッチインタラクション

運動視差とは頭部位置による視差を示し、頭部位置

による見え方の違いを表現する必要がある．図 3 に示すように，一般的な 3 次元カメラモデルではカメラ座標に対して固定された投影面を考えるが，カメラモデルの投影平面とディスプレイ平面を仮想的に一致させることで，移動するユーザの視点からディスプレイをのぞき窓として仮想空間を見るような状況が実現できる．

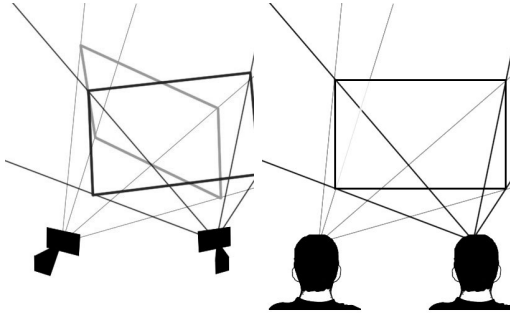


図 3 一般的なカメラモデルと運動視差カメラモデルの比較．
Fig. 3 Comparison of standard camera model and motion parallax-based camera model

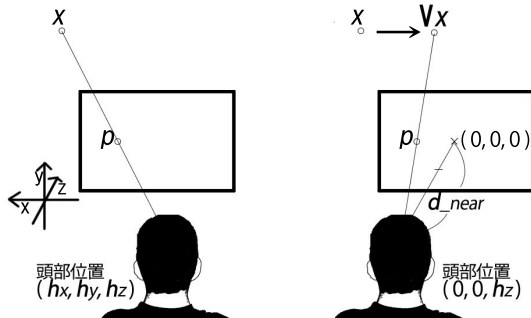


図 4 運動視差の幾何学モデル
Fig. 4 Motion parallax geometry

ここでは，上記のような運動視差を実現するためのカメラモデルについて説明する．視点を変更した場合，物体は投影平面上でずれて表示される．仮想空間中で投影平面を固定するために，このずれを正面から見た場合について求める．図 4 に示すように，頭部座標 (h_x, h_y, h_z) の視点から見たとき，3 次元空間中のある点 x が投影平面の点 p に射影されているとする．仮に Z 軸上から見たときにこの点 x が Vx に移動したと仮定すると，この点は投影平面上で点 p に射影される．したがって点 x が点 Vx にあれば，視点が Z 軸上から見た場合でも投影平面上の同じ点 p に射影される．

このときの運動視差の変換行列 V は，頭部座標を (h_x, h_y, h_z) ，頭部位置から投影平面までの距離を d_{near} として

$$V = \begin{bmatrix} 1 & 0 & h_x & h_x * d_{near} \\ 0 & 1 & h_y & h_y * d_{near} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

である．ここで 3 次元空間中の点が投影平面に射影される式は， q ， $M(q)$ および 3 軸移動パラメータ $[t_x, t_y, t_z]$ ，3 軸回転パラメータ $[q_x, q_y, q_z]$ を用いて

$$M(q) = T(t_x, t_y, t_z)Q(q_x, q_y, q_z)$$

となる． q はモデルビュー変換¹³⁾ に用いられる移動，回転の各パラメータであり， $M(q)$ はそのときのモデルビュー行列である．ただし，ここでは簡単のため，点 x は物体のローカル座標系ではなく世界座標系で記述されているとする．このときの移動行列 T は，

$$T(t_x, t_y, t_z) = \begin{bmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

またクォータニオン回転行列 Q は，

$$Q =$$

$$2 \begin{bmatrix} \frac{1}{2} - q_y^2 - q_z^2 & q_x q_y + q_w q_z & q_x q_z - q_w q_y & 0 \\ q_x q_y - q_w q_z & \frac{1}{2} - q_x^2 - q_z^2 & q_w q_x + q_y q_z & 0 \\ q_w q_y + q_x q_z & q_y q_z - q_w q_x & \frac{1}{2} - q_x^2 - q_y^2 & 0 \\ 0 & 0 & 0 & \frac{1}{2} \end{bmatrix}$$

となる．ただし，

$$|q| = \sqrt{q_w^2 + q_x^2 + q_y^2 + q_z^2} = 1$$

である．したがって運動視差が伴う環境下でのモデルビュー変換 $M'(q)$ は

$$M'(q) = V(h_x, h_y, d_{near})T(t_x, t_y, t_z)Q(q_x, q_y, q_z)$$

となる．この変換によりカメラモデルを変更せずに視点からの見え方の変化を表現する．

マルチタッチインタラクションを実現するためには，物体上のある点がどのように移動および回転をすれば現在のタッチ点に射影されるか計算するアルゴリズム³⁾が必要である．物体上の点 x はモデルビュー変換 M' および射影変換 P を行うことで，投影平面上の点 p に射影されるとすると，この変換を $s(x, q)$ として

$$p = s(x, q) = h(PM'(q)x)$$

となり，この一連の射影変換の過程に関してエネルギー関数 E を

$$E = \sum_i \|s(x_i, q) - p_i\|^2$$

と設定する．ここで i はタッチ点のインデックスを示

し, x_i は初期タッチ点の物体上の座標, p_i は操作終了時のディスプレイ座標である. 物体上の各タッチ点 x_i が q という移動および回転を行い投影平面に射影された点と現在のタッチ点 p_i の差の 2 乗の合計がエネルギー E である. このときのエネルギー E を最小にする移動および回転パラメータ q を Levenberg-Marquardt algorithm を用いて最適解を求める.

3.2 直感的なマルチタッチインタラクション

第 1 章で論じたように, このエネルギー最小化問題を解いただけでは, 図 1 のように 3 次元空間中のタッチ点のずれが生じる. 提案手法では, 操作中の物体を視差が生じない平面に移動することでこの問題を解決する. 運動視差には「仮想空間中の投影平面からの距離に応じて物体が移動したように投影平面に射影される」という特徴がある. したがって変換行列 V で使用される投影平面までの距離 d_{near} を, 物体上のタッチ点と同じ深さに設定することでタッチ点の深さで視差を最小化できる. タッチ点が複数ある場合は, これらのタッチ点の投影平面に対する深さの差を最小にするように投影平面を設定する.

また 3 点でタッチしたときそのタッチ点が 3 次元空間中の同じ奥行きにある場合, 奥行き方向の曖昧性が解決できないため, 図 2 のように意図しない方向に回転運動が生じる. エネルギー関数 E の最小化問題を求める場合, 通常は移動と回転のパラメータ q はタッチした時点で全てのパラメータを 0 に初期化する. しかし回転パラメータを 0 以外の値で初期化させることにより, その値の軸方向への回転が誘発される³⁾. これは Reisman らの文献でも議論されているが, どのような時にどのような値でパラメータを初期化すれば意図した方向への回転が誘発できるかという問題が依然として残っていた. 本研究では, これを頭部情報, とりわけ頭部位置の変位に注目することで解決する. 自然な回転方向は操作者の頭部情報と密接に関係しているため, 頭部情報により初期化する値を変化させることで, 自然な方向の回転運動が実現出来ると考える.

提案手法では, 図 5 に示すように頭部を負の方向に移動させているときには正の方向の回転が生じる方向に, 正の方向に移動させているときには負の方向の回転が生じる方向に値を設定する. すなわち, 頭部位置の変位が 0 のときは 0 で初期化をし, 頭部位置の変位が最大るとき回転パラメータが最大値を取るように値を対応付ける. 頭部位置の変位速度と回転パラメータを線形に対応付けすることにより, 問題なく回転が誘発されることを確認した.

以上のアルゴリズムを用いることで, 1 点タッチで

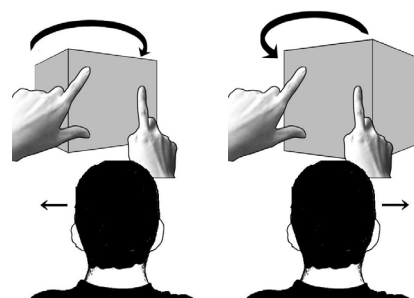


図 5 頭部情報を用いた回転.

Fig. 5 Rotation using head information

は XY 軸移動, 2 点タッチでは XYZ 軸移動および Z 軸回転, 3 点タッチでは任意の方向の移動および回転の操作が可能である. 4 点以上のタッチにおいても計算が可能であり, 同じアルゴリズムで全ての操作を実現する.

4. 実験

以上の提案システムに基づき, 評価実験を行った. ディスプレイは 3M 社の M2256PW を用いた. 静電容量方式を採用しており, 20 点までのタッチに対応している. またカメラ映像から頭部情報を取得するために Seeing Machines 社の FaceAPI を用いた. エネルギー最小化問題を解くための Levenberg-Marquardt algorithm は levmar: Levenberg-Marquardt nonlinear least squares algorithms in C/C++ を用いた. Intel Core i7 @ 3.40GHz processor, 12GB RAM の PC で実装を行った結果, 約 60 [fps] での動作を実現している.

4.1 実験設定

提案手法の操作精度, 操作時間および直感性を評価するために, 3 次元物体を空間中に配置する次のような実験を行った. まず各辺の長さが異なる 3 角形を用意し, 同じ形状の目標物体を移動および回転した状態で配置する. 被験者は視点変更をしながら回転および移動を行い位置合わせを行う.

本研究の物体の操作では, 1 本指で面内移動 (XY 軸移動), 2 本指で面内移動および奥行き方向の移動 (Z 軸移動) と回転 (Z 軸回転), 3 本指では面内移動と奥行き方向の移動および全ての方向の回転 (XYZ 軸回転) が可能である. したがって実験を分析するためは, 操作するために必要な指の本数ごとにタスクの設定を分け, 表 1 のように設定した. 目標物体がどのように

<http://solutions.3m.com/>

<http://www.seeingmachines.com/>

<http://www.ics.forth.gr/~lourakis/levmar/>

カテゴリ	タスク設定
カテゴリ 1	XY 軸移動
カテゴリ 2	XYZ 軸移動
カテゴリ 3	XYZ 軸移動 + Z 軸回転
カテゴリ 4	XYZ 軸移動 + X 軸回転
カテゴリ 5	XYZ 軸移動 + Y 軸回転
カテゴリ 6	XYZ 軸移動 + XZ 軸回転
カテゴリ 7	XYZ 軸移動 + YZ 軸回転
カテゴリ 8	XYZ 軸移動 + XY 軸回転
カテゴリ 9	XYZ 軸移動 + XYZ 軸回転

表 1 タスク設定ごとのカテゴリ。
Table 1 Categories of each tasks

移動および回転をして配置されるかは、あらかじめ各カテゴリに複数設定しておき、それらをランダムで出題することとした。

目標座標に対し誤差が一定以下になった場合物体の色が変化するため、それを目標に位置合わせを行う。制限時間中、被験者はエンターキー操作により、初期状態から何度でもやり直すことが出来る。物体の色が変化してから 3 秒後に次のタスクに移行するが、1 つのタスクにつき 60 秒の制限時間を設けており、時間になると強制的に次のタスクに移る。以上のような環境で、各タスクにつき位置合わせに要した時間と最終的な位置誤差を測定した。ここで言う誤差は、各頂点につき 3 角形の 1 つ分のずれを 1 として計算している。

実験は 7 人の被験者に対して行った。いずれもマルチタッチによる 3 次元 CG の操作は不慣れた被験者であった。まずは被験者に対し操作の説明をした後、実際に操作してもらい慣れたところで実験を開始した。比較対象としてキーボードの方向キーを操作することでカメラアングルを変えられる Reisman らの手法を用意し、本システムと同じデータセットを用いて実験した。

4.2 評価

実験結果を図 6、図 7 に示す。比較対象の Reisman らの手法による操作と本研究の提案操作につき、各カテゴリでのタスク終了までの平均時間およびタスク終了時の位置の平均誤差を表わす。

時間で比較した場合、カテゴリ 1, 2, 3, 4 ではほとんど変化がないが、カテゴリ 5, 6, 7, 8, 9 では本研究の操作の方が時間が短かった。本研究では覗き込むという人間の自然な動作を回転方向の決定に用いており、回転運動が必要な操作ではこれが有効に働いていると言える。カテゴリ 5 の時間の減少は特に顕著であり、本研究の操作では Y 軸回転時に左右から覗き込む動作が有効であると言える。対してカテゴリ 1、

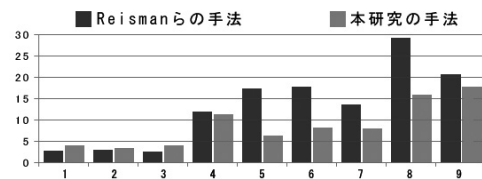


図 6 タスク完了までの所要時間の比較。
Fig. 6 Comparison of required times for task completion

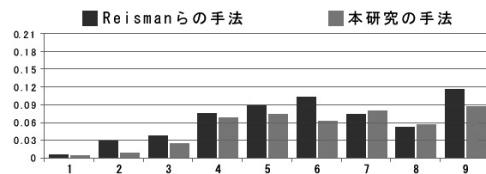


図 7 タスク完了時の位置誤差の比較。
Fig. 7 Comparison of position errors for task completion

2, 3 では、Reisman らの手法の操作と本研究の操作は共に単純であるため、違いが表れなかった。一方、回転運動を必要とするタスクであるにもかかわらず、カテゴリ 4 では平均時間が増加している。これは本研究の操作では X 軸方向への回転を誘発するために頭の上下運動が必要とされるが、本研究では実験を椅子に座った状態で行ったため、上下から覗き込むという動作が起こりづらかったと推測される。

位置の誤差で比較した場合、ほとんどのカテゴリで変わらないまたは誤差が少ないという結果が得られた。したがって Reisman らの手法と比べた場合、本研究の操作は正確性を損なうことなく、素早く物体の操作が行うことが出来ると言える。

さらに実験後に被験者に対し主観的評価の集計を行った。Reisman らの手法を用いた操作と本研究の操作ではどちらが好みであるか、それぞれの項目ごとに回答を得た。結果を図 8 に示す。XY 軸移動に関して Reisman らの手法が好ましいと答えた人数が多かったが、それ以外の項目では本研究の操作を選んだ被験者が多かった。総合的に判断した場合では、全ての被験者が本研究を選んだ。

以上から判断するに、既存研究である Reisman らの手法と比較した場合、提案システムの操作の方が任意の軸に対して自由な移動および回転が行いやすいことが確認できた。

5. 結論

本研究は、運動視差を用いた視点が自由に換えられる状況でマルチタッチインタラクションを行う場合、直感的なインタラクションを可能にする手法を提案し、

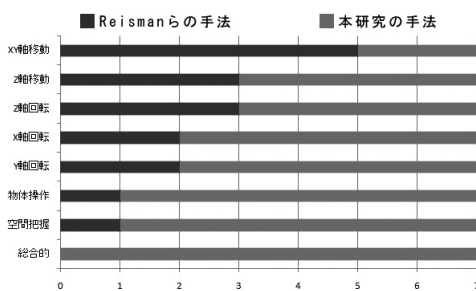


図 8 主観的な評価の比較 .

Fig. 8 Comparison of subjective evaluation

比較実験にてその有用性を示した .

運動視差による 3 次元空間中のタッチ点のずれが生じる問題では、投影平面をタッチ点に設定することでタッチ点のずれを最小化する解決を提案したが、この方法では依然として複数点でのタッチではタッチ点のずれが生じてしまう可能性がある . それにより意図しない物体の状態変化が起こることも考えられる . これを解決するためには、各タッチ点のずれを個別に最小化するようなアルゴリズムが必要である .

また回転運動が同一の方向のみ発生する問題では、頭部情報を回転パラメータの初期化の値に反映させる解決法を提案したが、頭部姿勢や物体の状態によりどのような値を設定すれば、より自然な回転が実現できるかについては更なる議論の余地がある . さらに頭を動かしていない状態では未だに回転の曖昧性は残っており、視線など頭部情報以外の情報を使用することが今後の課題として挙げられる .

参 考 文 献

- 1) M. Hancock, T.T. Cate, and S. Carpendale. "Sticky tools: Full 6DOF force-based interaction for multitouch tables." In *Proc. ITS*, pp. 145-152, 2009.
- 2) A. Martinet, G. Casiez, L. Grisoni. "The Design and Evaluation of 3D Positioning Techniques for Multi-touch Displays," In *Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology. VRST 2009*.
- 3) J.L. Reisman, P.L. Davidson, and J.Y. Han, "A screen-space formulation for 2d and 3d direct manipulation," In *UIST '09: Proceedings of the 22nd annual ACM symposium on User interface software and technology*, pp. 69-78, ACM, 2009.
- 4) A. Martinet, G. Casiez and L. Grisoni. "The Effect of DOF Separation in 3D Manipulation Tasks with Multi-touch Displays." In *Proceedings of VRST'10, the 17th ACM Symposium*

on Virtual Reality Software and Technology, 8 pages. ACM.

- 5) M. Hancock, O. Hilliges, C. Collins, D. Baur, and S. Carpendale. "Exploring Tangible and Direct Touch Interfaces for Manipulating 2D and 3D Information on a Digital Table." *ITS 09*
- 6) H. Benko, A.D. Wilson, and R. Balakrishnan. "Sphere: Multi-Touch Interactions on a Spherical Display." In *UIST '08: Proceedings of the 21st annual ACM symposium on User interface software and technology (2008)*, pp. 77-86.
- 7) J.B. de la Rivière, C. Kervégant, E. Orvain and N. Dittlo. "CubTile: a Multi-Touch Cubic Interface." In *VRST '08 Proceedings of the 2008 ACM symposium on Virtual reality software and technology*.
- 8) P.L. Davidson, J.Y. Han. "Extending 2D Object Arrangement with Pressure-Sensitive Layering Cues." In *UIST '08 Proceedings of the 21st annual ACM symposium on User interface software and technology*.
- 9) D. Valkov, F. Steinicke, G. Bruder, and K.H. Hinrichs. "2d touching of 3d stereoscopic objects." In *ACM Proceedings of CHI 2011 Conference on Human Factors in Computing Systems*, ACM, 2011.
- 10) F. Steinicke, K.H. Hinrichs, J. Schöning, A. Krüger. "MultiTouching 3D Data: Towards Direct Interaction in Stereoscopic Display Environments coupled with Mobile Devices." *Advanced Visual Interfaces (AVI) Workshop on Designing Multi-Touch Interaction Techniques for Coupled Public and Private Displays 2008*.
- 11) D. Valkov, F. Steinicke, G. Bruder, K.H. Hinrichs. "Traveling in 3D Virtual Environments with Foot Gestures and a Multi-Touch enabled WIM." In *Proceedings of Virtual Reality International Conference, VRIC 2010*.
- 12) J.B. de la Rivière, C. Kervégant, E. Orvain, N. Dittlo, and M. Courtois. "3d multitouch : when tactile tables meet immersive visualization technologies." In *Siggraph 2010 Emerging Technologies*, Siggraph 2010 Talk, 2010.
- 13) M. Gleicher and A. Witkin. "Through-the-Lens Camera Control." In *Proceedings of the 19th Annual Conference on Computer Graphics and Interactive Techniques J. J. Thomas, Ed. SIGGRAPH ' 92. ACM, New York, NY*, pp. 331-340. 1992.
- 14) M. Hachet, B. Bossavit, A.C. Cohé, J.B. de la Rivière. "Toucheo: Multitouch and Stereo Combined in a Seamless Workspace." *UIST 2011*.