

小袖屏風を対象とした インタラクティブ展示システムの開発

Development of an Interactive Exhibition System for Kosode Byobu

濱上知樹

HAMAGAMI Tomoki

- ①序論
- ②インタラクティブ展示システム
- ③鏡面球を用いた光源方向推定
- ④全形サンプル画像群の取得
- ⑤実装
- ⑥実験
- ⑦結論

【論文要旨】

近年の情報技術の進展に伴い、文化財のデジタルアーカイブ化が急速に進んでいる。また、その技術や利活用について多くの研究が行われている。デジタルアーカイブの研究は、人文系研究者の研究資料として重要な意味をもつことはもちろん、高度な展示技術への応用も期待されている。本研究においても、国立歴史民俗博物館が所蔵する小袖屏風を対象としたデジタルアーカイブ化と、これを利用した展示システムの構築を進めている。

その中で、実物を鑑賞した際に得られる感動や印象の再現を目的とし、インタラクティブ展示を提案する。インタラクティブ展示とは展示インタフェースと観察者との位置関係に応じて出力が変化する展示方式である。

頭の移動による試料の見え方の変化が光沢感の知覚を増大させるという結果が出ている。普段人間が質感を観察する際に、複数の角度から同じ対象を眺め、確認していることからわかるように、インタラクションを再現することは実際に物を見るという行為の再現につながるという。

本研究ではタブレット端末を対象とし傾きセンサを利用したインタラクティブ展示システムの開発を行ってきたため今回大型ディスプレイにおける全形のインタラクティブ展示システムの開発を目的とする。

インタラクティブ展示のレンダリングにおいては、ある程度平面の対象物を前提とし、様々な角度から光源を当て取得した画像群の線形補間により出力を行っている。しかし、データ取得のために多くの時間とコストがかかる点が問題である。今後アーカイブされる対象の増加を踏まえると、簡便な方法で、サンプル画像群を取得する必要がある。

そこで本研究では、手持ち光源により自由に光源を当て撮影した動画から、サンプル光源方向に紐づく画像群の抽出を行う。その際光源の方向を推定するため鏡面球をマーカーとして同時に撮影し、正射影を仮定することで画像座標のみから簡単な計算により推定を行う。そして、全形を復元するため試料の部分領域ごと得られたサンプル画像群をイメージモザイクにより結合する。イメージモザイク処理ではあらかじめ結合ずれが小さいと仮定し、Pyramid Blendingとグラフカットを用いたブレンディングを行う。

実験ではサンプルする光源間隔に対して誤差率約8%の精度で光源方向を取得することが可能であった。またグラフカットとピラミッドブレンディングを用いたイメージモザイク処理により従来に比べ大型ディスプレイによるインタラクティブ展示が可能となった。

【キーワード】 インタラクティブ、Pyramid Blending、レイトレース、BRDF、BTF

①……………序論

1.1 研究背景と目的

近年の情報技術の進展に伴い、文化財のデジタルアーカイブ化に注目が集まっている。デジタルアーカイブ化は文化財の保存や提供をすることができ、その技術や利活用について多くの研究が行われている [1] [2]。デジタルアーカイブの研究は、人文系研究者の研究資料として重要な意味をもち、さらに高度な展示技術への応用も期待されている。

高度展示においては、実物を鑑賞した際に得られる感動や印象を、いかにデータとして提示できるかが課題となる。従来研究として、高精細画像や、詳細な物理モデルによる自由視点、仮想環境下での再現手法が提案されている [3] [4]。しかし、単なる高精細画像では、対象のもつ物理的な応答は再現されない。また、物理的な再現を目指した手法では、データ取得のために多くの時間とコストがかかる点が問題である。今後アーカイブされる対象の増加を踏まえると、簡便な方法で、その場に「いる」「ある」といった存在感を再現できる展示方式が求められる。

本研究では、その場で見ているかのような印象再現を、「高リアリティ」と呼ぶ。類似の研究として、文献 [5] のように「写実リアル」と「感性リアル」に着目した研究がある。「写実リアル」とは、視覚的に正しい表現であり、「感性リアル」とは人の感性を刺激しより強い現実感を与えるような表現である。

本研究における「高リアリティ」とは「写実リアル」と「感性リアル」を組み合わせることを意味する。たとえば本研究が対象とする小袖屏風においては、高精細化という写実性だけではなく、光源や視点の位置によって変化する金糸の煌きや綸子の微細な陰影等の質感が存在感や印象に大きな意味をもつ。特に光沢感に関して、両眼と試料との位置関係による見え方の違いや頭の移動による試料の見え方の変化が光沢感の知覚を増大させるという結果が出ている [7] [8] [9]。普段人間が質感を観察する際に、複数の角度から同じ対象を眺め、確認していることからわかるように、両眼視やインタラク션을再現することは実際に物を見るという行為の再現につながるといえる。

そこで本研究では、写実性に加えインタラク션을与えることで高リアリティに質感を再現するようなインタラクティブ展示システムの開発を目的とする。

②……………インタラクティブ展示システム

本章では、高リアリティのためのインタラクティブ展示について述べる。光沢感に関して、頭の移動による試料の見え方の変化や両眼と試料との位置関係による見え方の違いが光沢感の知覚を増大させるという結果が出ている [7] [8] [9]。普段人間が質感を観察する際に、複数の角度から同じ対象を眺め、確認していることからわかるように、両眼視やインタラクçõesの再現が実際に物を見るという行為の再現につながるといえる。

そこで本研究では、写実性に加えインタラクçõesを与えることで高リアリティに質感を再現するようなインタラクティブ展示システムの開発を目的とする。

2.1 高リアリティ

1.1 で高リアリティを、その場で見ているかのような印象再現と定義した。そしてそのために画像の写実性に加えインタラクションを加えることで、高リアリティが想起すると仮定している。より一般的には実際に見る場合と、ディスプレイ上で見る場合との違いを補うことができれば、高リアリティに近づくと考えられる。

まず、ディスプレイは画素数以上の情報はもたないため、図 2.1 のように現実には知覚できる情報に比べ、ディスプレイは情報が欠損する方向に作用すると考えられる。そして、ディスプレイ上で見る場合に欠損し正確に伝達されない情報とは、少なからず以下の 4 つが挙げられる。ただし解像度については、写実性の高さに依存しているとする。

- ・ 距離情報
- ・ 相互作用（視点と物体の位置関係に応じた見え方の変化、自由視点性）
- ・ 縮尺
- ・ 色（光源の違いやディスプレイの反射特性を原因とする色の違い）

ここで人の視覚において、実際に物体を見る場合とディスプレイ上で見る場合との違いが上述した 4 要素の欠損と等しいのであれば、4 つの要素を補うことで高リアリティが想起されるといえる（図 2.2）。

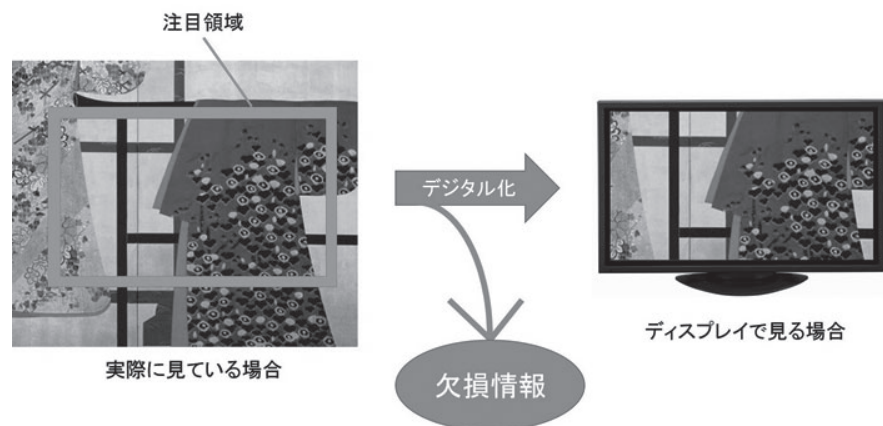


図 2.1 実世界内の現実物体とディスプレイ上の仮想物体の関係

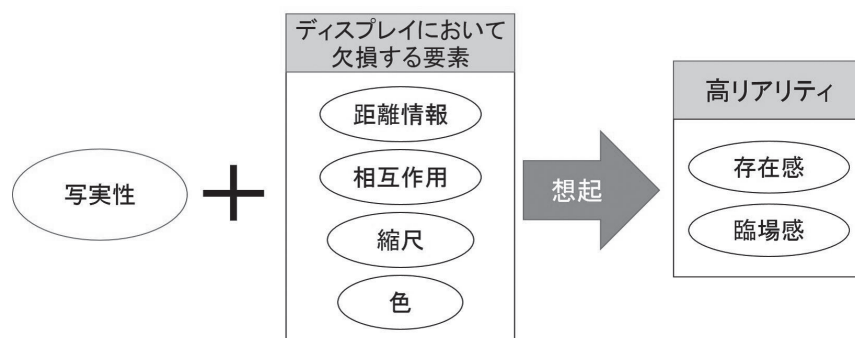


図 2.2 高リアリティのための要素と関係

2.2 インタラクティブ展示

インタラクティブ展示

インタラクティブ展示とは物体を見る際に起きる時間的な変化を観察者の行動に応じて再現する展示システムとして、本研究では定義する。具体例としては、図2.3の様にタブレット端末の傾きに応じた出力や、人間の目の位置に応じた出力を行う方式などが考えられる。先行研究ではタブレット端末において、様々な方向から試料に光源を当て撮影したサンプル画像群の線形補間により、傾きに応じた出力を行い、質感の分かりやすい展示が可能となっている[6]。そこで本研究では、大型ディスプレイを対象とし、図2.3右のような展示システムの開発を目指す。

インタラクティブ展示システムにおいて必要な要素は観察者とディスプレイの位置関係を取得するセンシング部分とその関係に応じて自由に画像を提供するレンダリング部分の2つに大別される(図2.4)。本研究では特にレンダリングに焦点を当てて研究を行っていく。センシングに比べレンダリングは小袖屏風という対象物に強く依存しており、小袖屏風特有の課題や効率的な手法が存在している可能性が高いからである。



図2.3 高リアリティのためのインタラクティブ展示

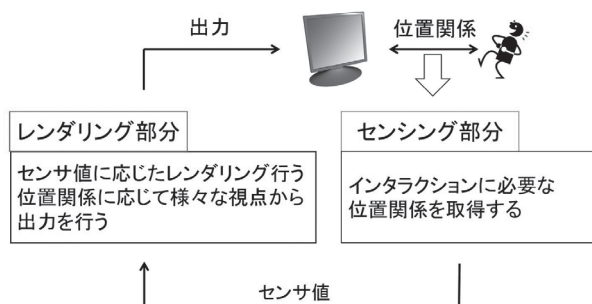


図2.4 インタラクティブ展示システムの概観

2.3 関連研究

本研究で提案するインタラクティブ展示システムにおいては観測者とディスプレイの位置関係に応じた視点から高精細な画像を提供することが必要となる。CG 技術においては写実的な画像を高速に生成することを中心的な目的として多くの描画アルゴリズムが提案されてきた。

特にフォトリアリスティック（写実性）の追求においては光の挙動をいかに再現するかということが重要であり、周囲の環境光、物体の形状による陰影といった光の挙動のモデル化をグローバルイルミネーションと呼ぶ。例として、レイトレーシングや IBL, PRT [10] といった手法があり、大域的な光の挙動をモデル化している。

一方、物体表面に光が照射し、物体の表面や内部における反射、屈折、散乱といった光の挙動のモデル化をローカルイルミネーションと呼ぶ。例として、BRDF や BSSRDF, BTF などがあり物体表面の局所的な光の挙動をモデル化している。

CG は従来、人の皮膚や布の様に微小な凹凸や色の変化をもつ面の表現には向いていなかったが、詳細な BRDF や BTF が実際にレンダリングへ適応できるようになりフォトリアリスティックに大きな効果をもたらした [11]。物体の質感を再現するためには、ローカルイルミネーションが重要であり、物体に適したモデルを利用する必要がある。

BRDF は一般的な反射をモデル化した双方向反射率分布関数であり、物体表面のある点 x における光の入射方向 w_i と視点方向 w_o を試料の垂直方向に対する極座標として変数にもった反射率分布関数である（図 2.5 左）。

微細な凹凸をもつような物体表面の BRDF を数式によりモデル化することは困難であったが、近年 BRDF を直接計測することで、反射特性を忠実に再現できるようになり、微細な質感の写実性に大きく貢献している。

また BTF はテクスチャの様にメゾスケールの光の振る舞いをモデル化した双方向テクスチャ関数であり、テクスチャの各テクセル上の BRDF が定義された関数である。

BTF よりテクスチャで全体を表現できる布等を写実的かつ効率的にレンダリングが可能である [9]。BTF は試料面を覆う半球状のあらゆる光源方向、視点方向の組み合わせに対し撮影を行い計測される。得られた画像群を BTF と呼ぶ場合もあるが、ここではこの画像群を BTF データと呼ぶ（図 2.5 右）。

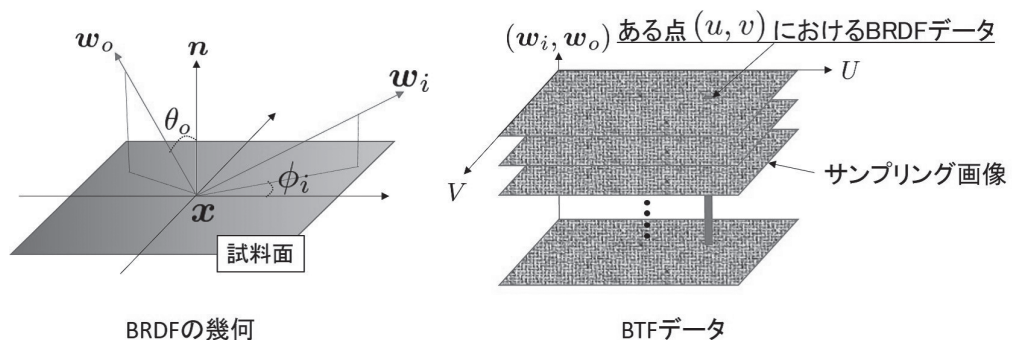


図 2.5 BRDF と BTF データの関係

BTFの利用例として能装束を対象とし金糸のサチュレーションを防ぐためHDRを利用し、視点の位置に応じた拡大縮小をスムーズに行うため解像度の異なるBTFレンダリング[4]やPRT[10]により、薪による動的な環境光を取り入れた舞台上での能装束レンダリングが実現されている[3]。

2.3.1 線形補間レンダリング

本研究では詳細な反射特性や3次元構造のモデリングを行うモデルベースな手法ではなく、試料を撮影した画像をそのままレンダリングするイメージベースな手法を用いる。固定視点下において様々な光源方向から撮影された画像群を取得し線形補間により、任意の光源方向から照らされた画像を生成する。

小袖屏風はある程度平面であり、3次元形状によるオクルージョンなどの見え方の変化が少ないため各光源方向でサンプリングされた画像群の線形補間でも十分写実的な自由視点画像を提供できると考えられる。またモデルの作りこみや、物理的なシミュレーションを含むレンダリングアルゴリズムが無いため実装が非常に簡易に行える。

サンプリングされた多方向照明画像群をサンプル画像群 $T = (T_0, T_1, \dots, T_N)$ とし、ある光源方向における画像を近傍のサンプル画像により線形補間で描画を行う。 T_i は光源方向 $w_i = (\theta_i, \phi_i)$ において撮影された画像を指す。任意の方向 w_* に対する出力を4近傍の $T_{n1}, T_{n2}, T_{n3}, T_{n4}$ の角度の比から線形補間を行い出力 \hat{T}_* を生成する(図2.6)。中間画像として ϕ の比から画像を $(T_{n1}, T_{n2}), (T_{n3}, T_{n4})$ の組

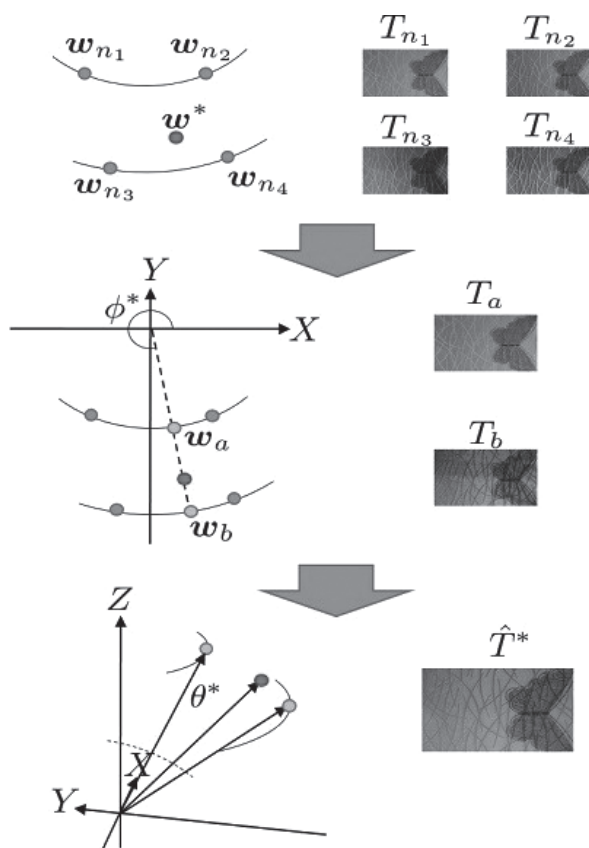


図 2.6 サンプル画像群の線形補間によるレンダリング

み合わせて合成したのち $T_a T_b$ を θ の比により合成する。この方式の場合、環境光を取り入れることや、サンプリング間隔の適切な設定が難しいというデメリットがある。しかし、モデルを持ち込まずに写実性の高い出力を自由な視点から取得可能である。またカメラを固定しているのは、今回ディスプレイが固定された状態での観察や、観察者が見る位置を変えずにディスプレイを変化させて観察する場合を想定しており、BRDFの光源方向と視方向を入れ替えても値が変化しないという性質から、固定視点下においてもレンダリングを行えるためである。

2.4 サンプル画像群取得のための課題と方針

2.3.1で述べたレンダリングではいかにして光源方向と紐付いたサンプル画像群を取得するかが課題となる。また、大型ディスプレイで表示するために全形の大規模なサンプル画像群を取得する必要がある。

先行研究では試料に対し垂直にカメラを固定し手持ち光源により自由に光を当て撮影した動画と、あらかじめ取得した光源方向が既知の5枚の画像を用いて、輝度値の類似性からSOMによりサンプル画像群の自動構築を試みている[6]。簡易的な撮影ではあるものの、試料の光源に対する変化が急峻な部分や変化の乏しい部分に関してはうまくサンプル画像群を構築できていない。そのため、金糸や織りの細かい布地領域を含む場合でも試料に対する光源方向を安定的かつ位相関係が正しくなるように取得する必要がある。

そこで、動画撮影時に光源方向を一意に計算できるようなマーカの撮影を同時に行い、サンプル画像群の取得を行う。本研究ではマーカとして鏡面球を利用し、生成した画像に対する正射影を仮定することで従来に比べ簡単な計算により光源の方向を推定する。その後得られた部分領域のサンプル画像群をそれぞれつなぎ合わせ、最終的な全形のサンプル画像群を取得する。

③……………鏡面球を用いた光源方向推定

本章では動画像における光源推定について述べる。2.3.1で述べたサンプル画像群は撮影された際の光源の方向と紐付いた画像群であり、これはBTFデータと同様である。本研究では特殊な計測装置を用いずに簡易的にサンプル画像群を取得するために、手持ち光源により自由に試料を照射した動画を用いる。

動画の各フレームにおいて光源方向を推定し、必要なサンプリング方向のフレームを抽出することでサンプル画像群を生成する。これは画像による光源環境推定問題に帰着する。物体の形状による陰影を用いた手法、物体の反射特性を用いた手法の主に2つのアプローチがある。今回対象としている小袖屏風はほぼ平面であり、反射特性も未知であるため陰影から直接光源方向を推定することは難しい。

よって、光源方向が一意に定まるようなマーカを同時に撮影することで各フレームの光源方向を取得する方法を利用する。マーカを用いた光源環境推定手法として鏡面球を利用する例[12]が多く、鏡面球とカメラの位置関係、鏡面が作り出すハイライトから鏡面球周囲の光源環境推定を行っている。

本研究においても鏡面球を用いて光源方向の推定を行う。その際、計測環境として、試料がカメラの光軸とほぼ垂直であること、鏡面球と試料面に対しカメラが十分に離れていることを前提とし

画像面に対する正射影を仮定し光源方向を求める。これにより、カメラと鏡面球の実座標における位置関係は必要なく、画像中の鏡面球の中心座標と半径のみで計算を行うことができる。

3.1 鏡面ハイライトの検出

鏡面ハイライトは、鏡面反射により光源がそのまま写しだされた領域であり、鏡面球において輝度値が最大となる領域である。実装においては、ノイズ除去のために Gaussian フィルタにより平滑化を行い、マーカ領域内で最大輝度値を取る各画素の座標の重心をハイライトの中心 h'' とした。

3.2 鏡面球中心における光源方向の取得

X-Y 平面上に試料面が存在し、カメラが Z 軸上に存在しているとする。また、点 P' に対する光源 I' の方向 w' が求めたい値であり、この点 P' 付近においては光源が平面光であるとする。

今回、カメラの画像が正射影により投影されていると仮定すると、Z 軸方向とハイライト方向のなす角 θ_t は図 3.1 の様に、 I' 方向に対し 1/2 倍の値をとることが分かる。そして簡単に取得可能な画像座標上の $\tilde{w}' = (\tilde{\theta}', \tilde{\phi}')$ を $w' = \tilde{w}'$ として次式で計算を行う。

$$\tilde{\theta}' = 2 \sin^{-1} \left(\frac{|h'' - P''|}{r''} \right) \quad (3.1)$$

$$\tilde{\phi}' = \begin{cases} \cos^{-1} \left(\frac{h''_x - P''_x}{|h'' - P''|} \right) & (h''_y - P''_y > 0) \\ 2\pi - \cos^{-1} \left(\frac{h''_x - P''_x}{|h'' - P''|} \right) & (otherwise) \end{cases} \quad (3.2)$$

以上の様に、画像面が正射影により投影されているとすれば、光源方向は簡単に計算を行えるが、理想的な投影モデルである透視投影においてどのような誤差が発生するのかを 3.3 で述べる。

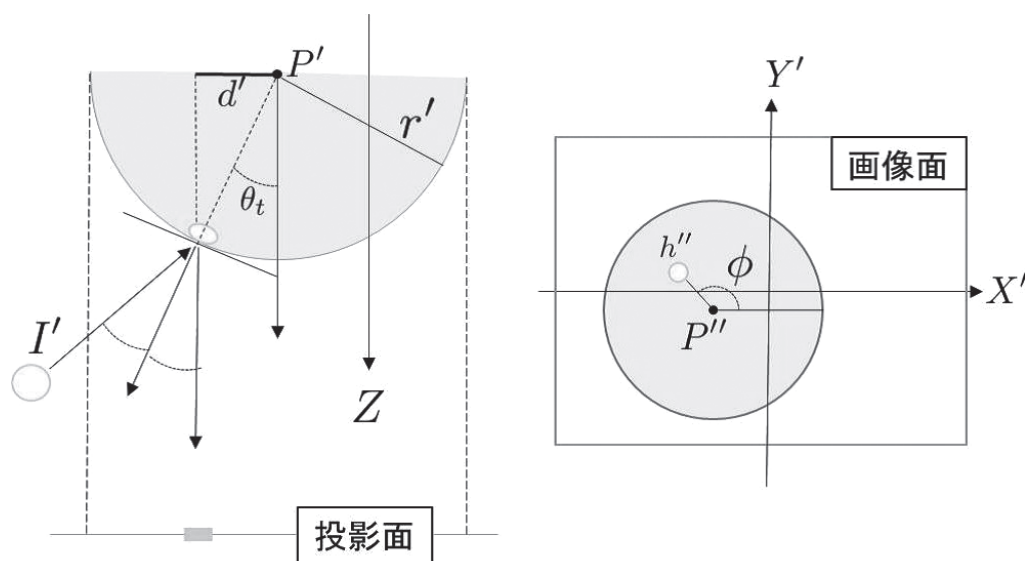


図 3.1 正射影におけるハイライトの位置と光源方向

3.3 正射影を仮定した計算誤差

ある光源位置 $I = (x_i, y_i, z_i)$ ，半径 r の鏡面球の位置 $P = (x_p, y_p, z_p)$ ，カメラ位置 $C = (0, 0, 0)$ が与えられたとき，3.2 の式 3.1，3.2 で計算される値を求めることを考える（図 3.2）。ただしカメラは理想的なピンホールカメラとしレンズの収差は考慮しない。

点 P から球面上のハイライト位置 h へ向かう単位ベクトルを $\mathbf{n} = (i, j, k)$ とすると \overrightarrow{PC} と \overrightarrow{PI} は \mathbf{n} に対してなす角が等しく，3つのベクトルは同じ平面上にあるため $\mathbf{t} = \overrightarrow{PI} / |\overrightarrow{PI}|$ ， $\mathbf{u} = \overrightarrow{PC} / |\overrightarrow{PC}|$ とおくと \mathbf{n} は次式の様になる。

$$\begin{aligned} \mathbf{n} &= \frac{\mathbf{t} + \mathbf{u}}{L_t} \\ &= \frac{1}{L_t L_i L_c} \begin{pmatrix} L_c x - (L_c + L_i) x_p \\ L_c y - (L_c + L_i) y_p \\ L_c z - (L_c + L_i) z_p \end{pmatrix} \end{aligned} \quad (3.3)$$

この時 $L_t = |\mathbf{t} + \mathbf{u}|$ ， $L_i = |\overrightarrow{PI}|$ ， $L_c = |\overrightarrow{PC}|$ である。 \mathbf{n} を用いて h は次式の様になる。

$$\begin{aligned} h &= P + r\mathbf{n} \\ &= \frac{1}{L_t L_i L_c} \begin{pmatrix} L_h x_p - r L_c x \\ L_h y_p - r L_c y \\ L_h z_p - r L_c z \end{pmatrix} \end{aligned} \quad (3.4)$$

この時 $L_h = L_t L_c L_i + r(L_c + L_i)$ である。理想的なピンホールカメラモデルの透視投影により h, P を投影面に写したものを h_p, P_p とすると，次式の様になる。

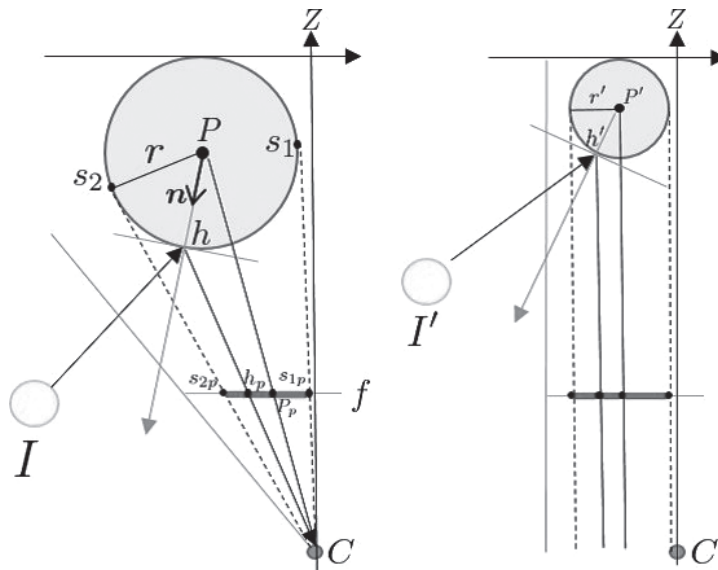


図 3.2 透視投影と正射影

$$h_p = \frac{f}{L_h z_p - r L_c z} \begin{pmatrix} L_h x_p - r L_c x \\ L_h x_y - r L_c y \\ L_h x_z - r L_c z \end{pmatrix} \quad (3.5)$$

$$Pp = \frac{f}{z_p} \begin{pmatrix} x_p \\ y_p \\ z_p \end{pmatrix} \quad (3.6)$$

また r' は $|s_{1p} - s_2|/2$ として求めると、次式のようになる。

$$r' = \frac{f(\tan(\theta_p + \theta_{sp}) - \tan(\theta_p - \theta_{sp}))}{2} \quad (3.7)$$

$$\begin{aligned} \theta_p &= \tan^{-1} \left(\frac{\sqrt{x_p^2 + y_p^2}}{z_p} \right) \\ \theta_{sp} &= \sin^{-1} \left(\frac{r}{\sqrt{x_p^2 + y_p^2 + z_p^2}} \right) \end{aligned} \quad (3.8)$$

式 3.1, 3.2 を h_p, Pp, r' により求めた値が \mathbf{w}' となる。本手法では h'', P'', r'' を利用するため画像座標上への変換にカメラの内部パラメータが必要になる。ここでは歪みの無いスケーリングのみの変換とすると h'', r'', P'' はそれぞれ次式の様になる。

$$h'' = \begin{pmatrix} f_w h_{px} + c_x \\ f_h h_{py} + c_y \end{pmatrix} \quad (3.9)$$

$$P'' = \begin{pmatrix} f_w P_{px} + c_x \\ f_h P_{py} + c_y \end{pmatrix} \quad (3.10)$$

$$r'' = \frac{(f_w + f_h)r'}{2} \quad (3.11)$$

ここで $f_w f_h$ はそれぞれ実座標を画像座標に変換する係数であり、 c_x, c_y 画像中心の座標である。これらは焦点距離 f 、撮像素子サイズ (W, H)、画素数 (N_w, N_h) を用いてあらかじめ計算可能な値である。

式 3.1, 3.2 により求めた光源方向と実際の光源方向 \mathbf{w} との差が、3.2 の光源方向の計算において想定される誤差である。3.3.1 ではその誤差をグラフにより確認し、マーカの配置と試料中心に対する光源方向の計算について述べる。

3.3.1 4点マーカの中央方向の計算

マーカは試料面と同時に撮影を行うが、撮影される試料領域はなるべく大きくすることが望ましく、また収差の小さい画像中心付近を利用したいため、マーカはなるべく画像中の4隅に収まるよう配置し撮影を行いたい。

その際、3.2で述べた方法はマーカの中心における光源方向を計算しているため、画像中央の試料の方向としては扱いにくい。よって4隅に配置した各マーカの光源の位置を求め、各方向から画像中央の光源方向を計算することを考える。

図3.3の様に4つのマーカから得られた光源方向を w_1, w_2, w_3, w_4 とし、4つのマーカのワールド座標系における位置を P_1, P_2, P_3, P_4 とする。

シミュレーションとして3.3の方法により誤差の理論値を求めグラフにし確認を行う。原点を O とした座標系 X - Y - Z においてカメラの位置 $C = (0, 0, 1000)$ 、各鏡面球の位置を $w_{1...4} = (110, 170, 0), (-110, 170, 0), (-110, -170, 0), (110, -170, 0)$ とした時に原点の光源方向 w_c と各鏡面球方向 w の誤差を ϕ, θ についてそれぞれプロットしたグラフが図3.4, 3.5である。光源は O を中心とした半径1,000の Z 軸正方向の半球上から原点に照射される点光源とする。

図3.4, 3.5から、原点 O に対して対象の位置に存在する P_1, P_3 と P_2, P_4 が ϕ, θ 共に絶対値がほぼ等しくかつ正負の反転した様な値をとることが分かる。そこで、4隅におかれた各鏡面球に対する光源方向の平均値 w_c を原点である試料中心に対する光源方向 w として計算を行うと、図3.6の様に誤差が打ち消しあい、推定したい w との誤差が小さくなっている。

図3.7は4点の鏡面球の推定値の平均値 w_c と実光源方向 w とのなす角のグラフである。シミュレーションのカメラと鏡面球の位置関係においては、ほとんどの光源方向に対し誤差となる角度が0.1未満に収まることがわかる。

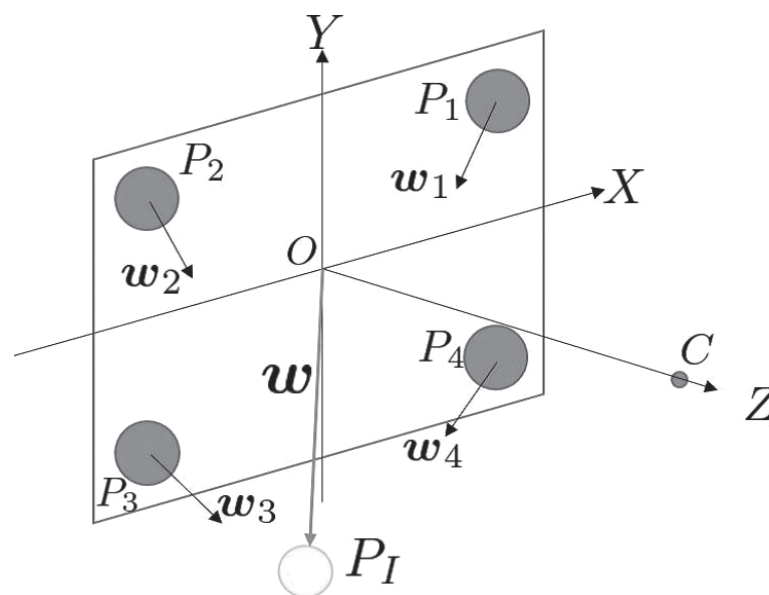


図3.3 4点の鏡面球マーカと試料中心の光源方向

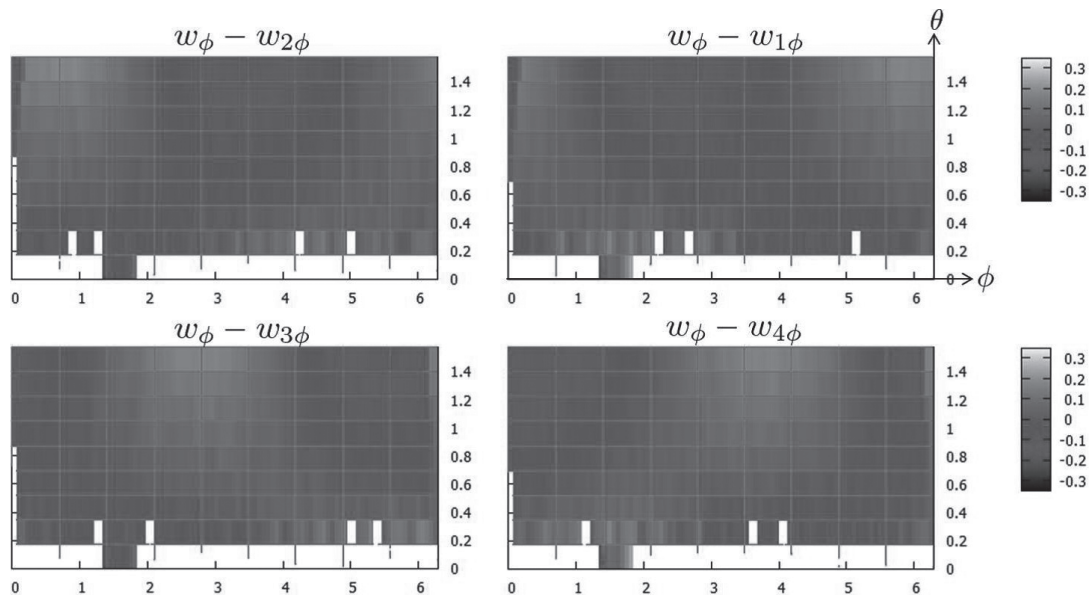


図 3.4 4 点の鏡面球と w との ϕ 方向における推定誤差

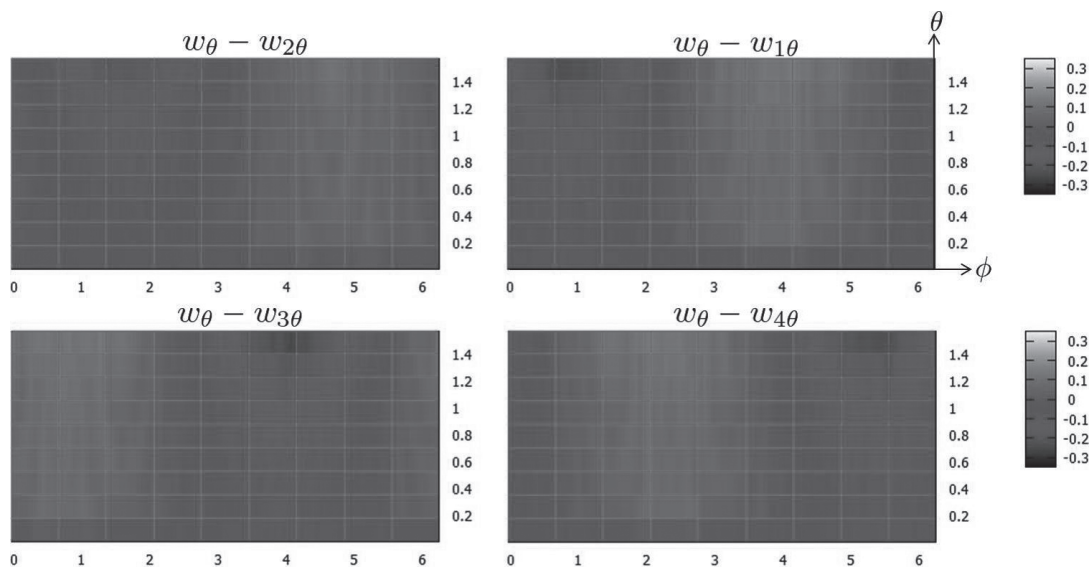


図 3.5 4 点の鏡面球と w との θ 方向における推定誤差

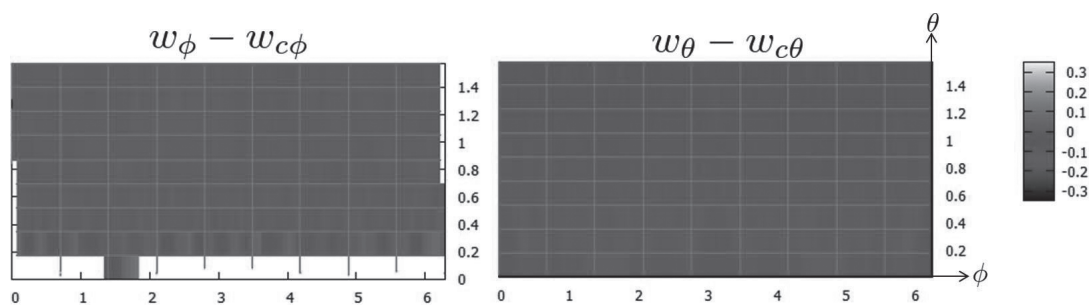


図 3.6 4 点の鏡面球の平均値と w との ϕ , θ 方向における推定誤差

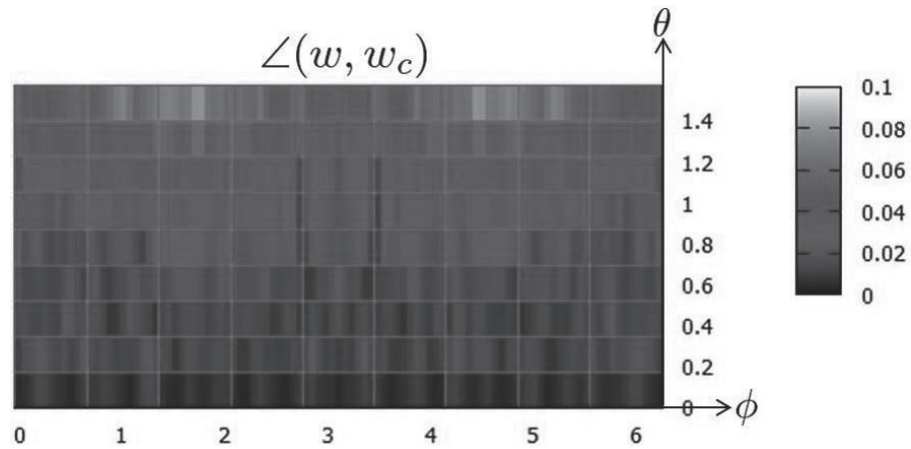


図 3.7 4 点の鏡面球の推定値平均と試料中心の光源方向の誤差

3.3.2 鏡面球とカメラの配置と誤差

一般的に正射影はカメラと対象物の距離が十分離れている場合、対象物の凹凸が小さい場合に良い近似となる。一方で、カメラがあまりに離れすぎると計算に整数値をとる画像座標を用いているため、量子化誤差が大きくなる。またカメラの光軸に対し鏡面球が垂直方向に離れすぎの場合、正射影の透視投影としての近似精度は悪化する。これらの関係性をグラフにより図示し、鏡面球とカメラの位置と推定誤差の関係を確認する。

まず、カメラの位置を $C = (0, 0, 1000)$ として固定し、4 点の鏡面球の光軸との垂直方向の距離とカメラと原点との距離の比と、その際の最大誤差を表したグラフが図 3.8 である。ただし最大誤差は ϕ, θ をそれぞれ $\pi/180$ ずつ変化させ推定した際の最大誤差とし、鏡面球の位置関係は 3.3.1 のシミュレーションと同様とし、各鏡面球の原点に対する距離だけを変化させている。

図 3.8 から d/z_p がおよそ 0.25 程度であれば最大でも誤差が 0.1 に収まる。特に 0.14 ~ 0.24 程度の精度が良いため、誤差の打ち消しあい最も良く表れているといえる。3.3.1 のシミュレーションにおける d/z_p は約 0.2 であり、位置関係としては理論上誤差の小さい比である。

そこで、シミュレーション同様にカメラと鏡面球を固定し、鏡面球の半径を変化させ最大推定誤差を確認したグラフが図 3.9 である。定性的には、 r/z_p が小さすぎる場合、量子化誤差の影響が大

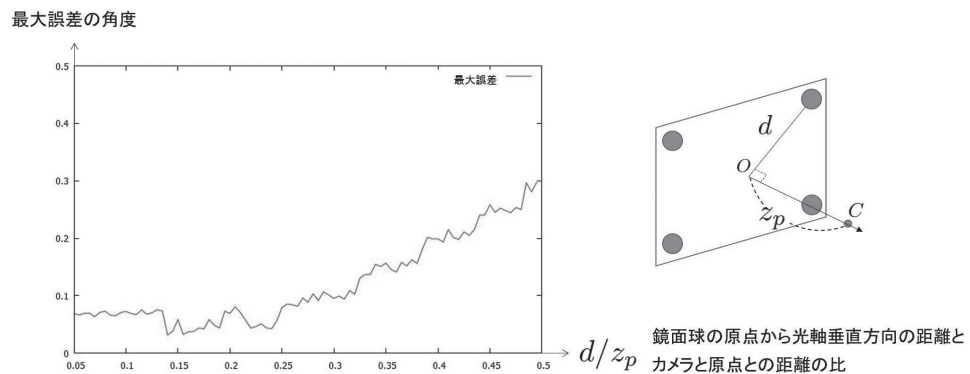


図 3.8 原点に対するカメラと鏡面球の距離の比と最大推定誤差

きくなり、大きすぎる場合、正射影の近似精度の悪化の影響が大きくなるといえる。そのため下に凸の曲線になると想定されるが、図3.9から分かるように、下に凸のグラフとなっており、 r/z_p が約0.005～0.05の範囲であれば最大でも誤差が0.1程度に収まる。シミュレーションの設定では r/z_p が0.0065となっている。鏡面球はなるべく小さい方が試料をより大きく撮影可能であるが、必要となる精度とのバランスを考慮して設定する必要がある。

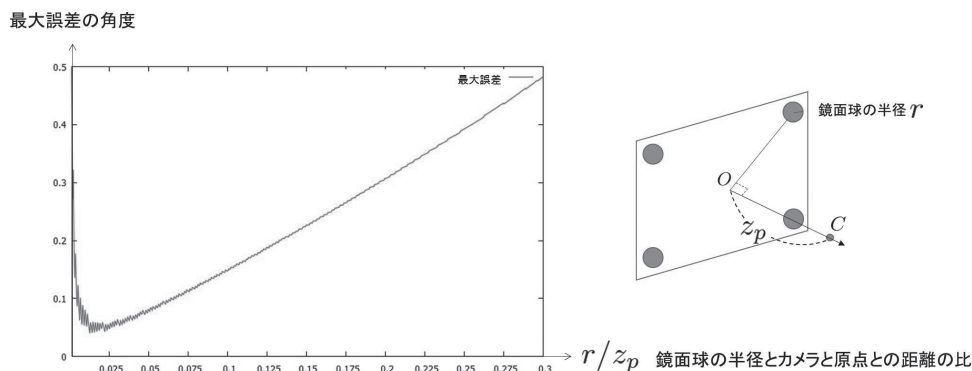


図3.9 原点とカメラの距離—鏡面球の半径比と最大推定誤差

3.4 推定精度の確認実験

実験として、 $1,920 \times 1,080$ の解像度の動画において4隅に鏡面球を配置し撮影を行う。実際の光源方向と本手法で求めた推定値の比較を行った。今回鏡面球として直径13[mm]の金属球を用意し、 249.6×140.4 [mm] の長方形に収まるよう4隅に4つの球を配置し光源推定マーカ作成した。撮影時はマーカの4点ができるべく画像面の4隅に収まるよう撮影を行う（図3.10左）。カメラと原点との距離は約725[mm]である。そのため3.3.2で述べた d/z_p が約0.2、 r/z_p が約0.009となっている。

4点のマーカの中央にもう一つマーカを置き、正解方向、中央の1点での計測値、4点での計測値をそれぞれ求め比較を行った。撮影用のカメラにはCanon社のEOS 6Dを用いた。光源は市販のLEDライトを用いており、色温度が約5,600 [K]、中心照度が約2,100 [Lx/50cm] 照射角約40度のライトである。

実験データは図3.10右の様に台の上の半円状ライトを等間隔に動かしていき計15点でサンプリングを行った。



図3.10 マーカの配置と実験データの撮影環境

3.4.1 結果・考察

実験データの15点で光源を推定した結果が図3.11である。グラフを見ると各サンプルに対し誤差が似たように分布していることが分かる。理論上どちらも誤差が非常に小さいため計測上の誤差やハイライトの検出誤差が同様に反映されているといえる。

0, 7, 14番目のサンプルにおいて特に誤差が大きく出ている。サンプル7においては図3.12を見ると鏡面球の下部が他のサンプルと比べ白くなっている。これは各機器の下にある白い台の反射光が正面になり強くなったため鏡面球の下に写りこみハイライトがずれたと考えられる。

サンプル0, 14に関しては、光が鏡面球のほぼ真横からあたっている状態である。その場合ハイライトの中心は見えている半球の縁（図3.12における赤い円上）に存在するはずだが最大輝度値の重心としていることからハイライト位置が中心方向にずれ、結果として誤差が生じていると考えられる。

実測値（正解値）との平均誤差は約2.5度、計測誤差を無視した場合の理論値との平均誤差は約0.9度であった。画像群のサンプル間の角度は約30度で取得しているため、サンプル角度に対する誤差率としては8%程度となる。

角度の誤差がインタフェースの観察者に対しどの程度の影響を与えるかは実験的に検証し精度目標を定める必要があるが、今回はサンプル角度に対し位相関係のある程度保った推定が行えれば十分とし本光源推定法を利用する。

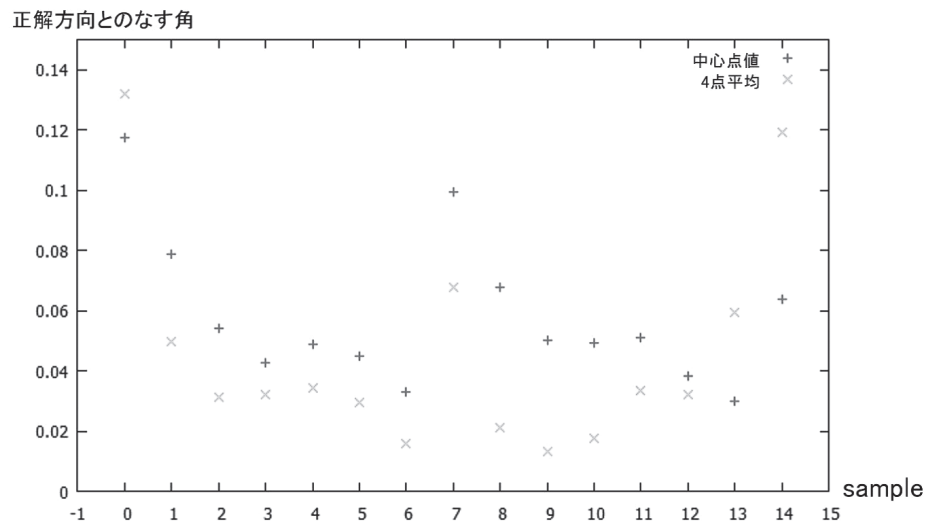


図 3.11 サンプル点に対する推定値と正解値のなす角

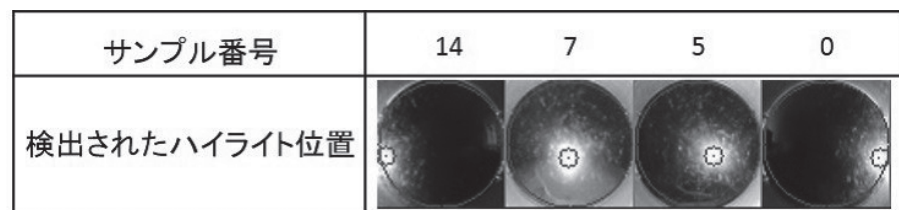


図 3.12 検出されたハイライト位置の例

3.4.2 サンプル画像群の抽出と確認

実際に、本手法により取得したサンプル画像群により〔6〕同様にタブレット端末におけるインタラクティブ展示を行ったところ良好な出力が得られた。本研究で目的としている大型ディスプレイによるインタラクティブ展示を行うためには、より大きな領域のサンプル画像群を取得する必要がある。単純にはカメラを対象から離しより多くの領域を画角に収め、それに合わせて鏡面球の半径を定めればよい。その際ディスプレイにおいて実寸表示かそれ以上の dpi となるようなカメラの位置とすれば十分であるとする。

そこで本研究の実験環境でなるべく大きな領域の計測をするため、対象と約 2,150mm 離し半径 20mm の鏡面球を用いることで約 700 × 400mm の領域を撮影し、サンプル画像の取得を行った。この場合、画像の実寸大における dpi がディスプレイの 1.5 倍程度となる。得られたサンプル画像を用いて 249.6 × 140.4mm の際のデータ同様タブレット端末一傾きセンサ、また大型ディスプレイ-Kinect (5.1) を用いたインタラクティブ展示を行ったところ良好な出力が得られた。そのため約 700 × 400mm 程度までの領域であれば、サンプル画像群を取得可能であるといえる。ただし実際にレンダリングに利用できる領域は鏡面球の分小さくなる。

3.5 全形サンプル画像取得のための課題

本手法で展示対象である小袖屏風を一度に撮影しきるのは難しいため、撮影領域を撮影可能な大きさに区切りそれらをサンプルした光源方向同士つなぎ合わせるか（図 3.13）、または CG の様に、一部の領域で取得した動画・サンプル画像群から他領域へ適用可能な BRDF データを抽出し、動画に比べ高解像度の静止画により取得した全形画像において、類似した領域または画素に BTF データ・BRDF データをマッピングする方法が考えられる（図 3.14）。

全形を分割して計測し光源各方向において画像をつなぎ合わせる

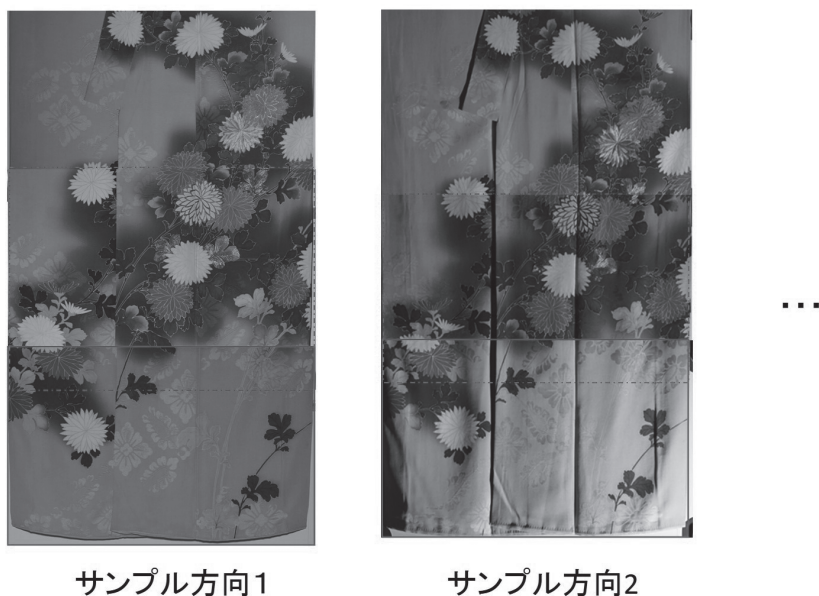


図 3.13 光源方向ごとの画像結合

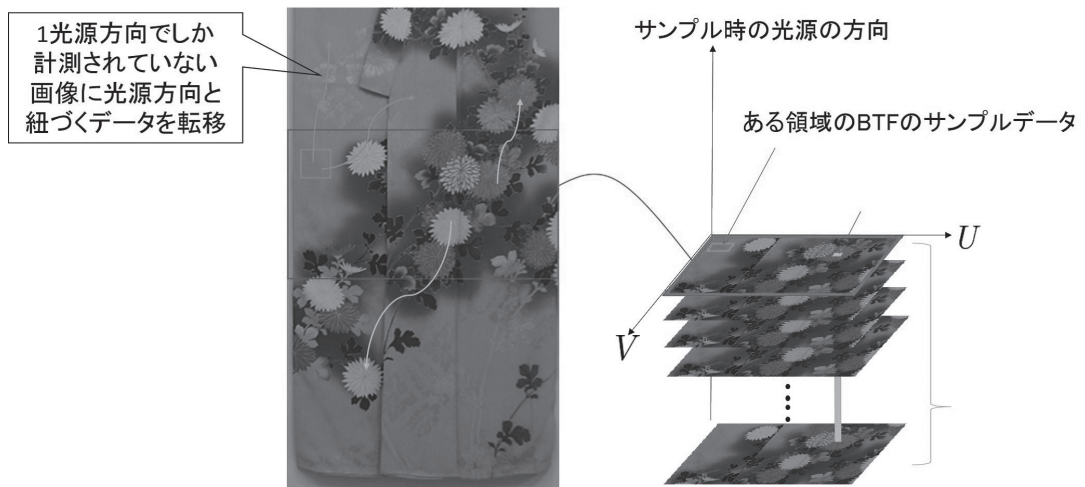


図 3.14 他領域へのサンプルデータの転移

後者の方法の場合、計測の手間は省かれる一方で、手持ち光源で取得した画像群は光源の照度が均一でなく、対象の形状モデル、BRDF モデルもないため一方向のサンプルから転移先を推定しその領域になじむように転移するのは難しいと考えられる。また小袖屏風自体もシミやほつれ大小さまざまな文様があることから、テクスチャを用いて描画すること自体が難しい、つまり全計測が必要となる場合も存在する。

そこで、本研究では可能な限り大きな領域で全形を分割しそれぞれ鏡面球を用いた計測を行い、画像群をつなぎ合わせることで全形のサンプル画像群を取得する。その際、計測回数を削減するためにもなるべく計測領域同士の重なりを小さくすることが望ましいが、オーバーラップする領域が小さくなるほど、画像のつなぎ合わせは難しくなる。

また、手持ち光源で撮影されたサンプル画像群は特に大きな領域になるほど画像間・画像内で光源の強度が均一でないため、つなぎ合わせた画像に明暗の境界など違和感の生じる可能性が高いと考えられる。そのため、重なり的小さい画像同士においても安定して結合が可能であり、各サンプル画像の光源方向のずれを補正する手法が必要であるといえる。

④……………全形サンプル画像群の取得

本章では小袖屏風に対し部分的に計測したサンプル画像群から全形のサンプル画像群の取得方法について述べる。一般的に複数枚の画像または動画から一枚の自然な画像を生成する処理をイメージモザイク [16] と呼ぶ。本研究においてもイメージモザイクにより、全形のサンプル画像群の取得を行う。以降、イメージモザイクについて述べ本研究の課題に沿った手法について述べる。

4.1 イメージモザイク

イメージモザイクは画像群を一枚の自然な画像にするために結合する処理である。そのため、画像群内の各画像は少なくとも一枚の他の画像と重なり合う領域をもつ必要があり、それらの結合領

域が可能な限り等しくなるように重ねる。

一般的にイメージモザイクで必要になる処理は主に以下の三つであり、図4.1のような流れで処理が行われる。以降順に説明を行う。

- ・対応点の取得
- ・画像の結合
- ・ブレンディング

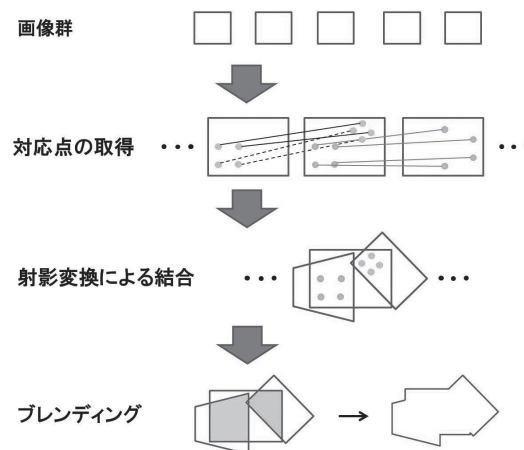


図4.1 イメージモザイクの処理概要

4.1.1 対応点の取得

画像処理の観点から画像の位置関係を求めるには、画像間で対応点を取得する必要がある。一般的に画像間の対応点を求める場合、各画像から特徴点を検出し、特徴点に紐づく特徴量を比較することでマッチングを行う。例えばSIFT特徴量は、画像の勾配情報からスケールに応じた特徴的な点を検出し、その近傍の勾配情報に基づき128次元のベクトルとして記述することで画像間のスケール、回転の変化にロバストな対応点の取得を行えるため、対応点取得法として広く用いられる。

また、動画の様に隣接する各画像間の変動が小さい画像群を利用する場合、高速・高精度に対応点を取得する手法としてKLT法がある[13][14][15]。

KLT法では特徴点として固有値を用いた画像中の特徴点の検出と、その近傍の小領域内で特徴点の画素値が大きく変化しないという制約から探索を行い画素値の二乗和が最小となる点を対応点として検出している。

4.1.2 画像の結合

求めた対応点から画像を結合に必要なパラメータを求める。画像を結合する最も単純な方法は、画像を平行移動して並べる方法であるが、この場合カメラの回転や傾きによるずれは修正できない。よって、撮影時のカメラの移動方向や向きの違いなどを考慮するために、アフィン変換、平面射影変換などの結合方法が利用される。これらの方法はカメラの撮影状況（位置や向き）の違いにより発生する画像結合時のずれを小さくすることができる。撮影状況の違いとそれを考慮した画像結合方法の関係を図4.2に示す。

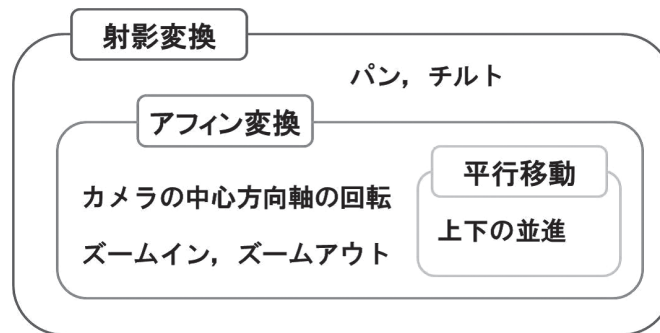


図 4.2 画像結合方法と撮影方法の関係

平面射影変換は変換後に対応点が最も合うように平面に対するカメラの外部的な変化を全て考慮し結合が行えるため、イメージモザイクにおいて広く利用されている。

しかし、今回の様に撮影条件として試料に対しカメラがほぼ垂直である、つまりパン・チルトの変化がないためアフィン変換でも十分にずれの小さい結合ができる可能性がある。また、平面射影変換は画像の変形が大きく、誤差が基準画像から離れるほど大きくなるため、重なり領域自体のずれが小さくても全体としてのずれが平行移動やアフィン変換に比べ大きくなる場合もあると考えられる。そのため、高精度の対応点取得とブレンディングにおけるずれの解消が求められる。

4.1.3 ブレンディング

一般的にブレンディングとは、複数の画像を合成する際に目的の画像になるよう、画素値を決定する処理である。

イメージモザイクにおいては画像間の変換が求めたのち重なり合う部分のつながり目や明るさなどを滑らかにする処理である。以下、代表的な手法について述べる。

α Blending 0～1の α 値により新規画像の画素値と元画像の画素値を $\alpha:1-\alpha$ の比率でブレンドを行う。 α 値を一定ではなく、重なる領域内で元画像と距離に近いほど値を減衰させるようにマスクを作ることによって、十分に自然な出力を得ることも可能である。一方で、エッジの強い領域においては、両方のエッジが合成時に残ってしまうゴースト（図 4.3）と呼ばれる現象が起きてしまう。



図 4.3 ゴーストの発生例

Pyramid Blending (PB) [17] 画像の低周波成分から高周波のエッジ成分に分割した画像ピラミッドにおいて、低周波の画像は緩やかに α 値を変更し、エッジの強い高周波成分ほど急峻に α 値を変更しブレンディングを行う。全体としての輝度値を滑らかにつなぎつつもゴーストの発生を抑えるようなブレンディングを行える。しかし、高周波成分を急峻につないでいるため、対応点自体にずれが発生している場合には、エッジ上にもそのずれが反映されてしまう。そのため、画像を重ね合わせ際のピラミッドの各レイヤの境界の設定方法が重要となる。

Poisson Image Editing (PIE) [18] 新規画像を元画像のある閉領域に重ねる際、その領域内の勾配と新規画像の画素値を用いてポワソン方程式を解くことで、新規画像のテクスチャを保持したまま領域の勾配に沿ったブレンディングを行える。PIE ではゴーストが起きにくく、つなぎ目の分かりにくい自然な合成が行えるため、テクスチャ合成やオブジェクト合成に広く利用されている。一方で元画像において周辺の画素値が存在する閉領域が必要となる。

4.2 本システムのイメージモザイク処理

3.5 で述べたように計測回数を削減するためにも部分領域同士の重なりはなるべく小さい方が望ましい。一方で重なり領域が小さい場合、画像全体で得られる特徴点に対し正解の対応点が少ないため誤対応の増加が問題として考えられる。そこで、計測した各画像の大まかな位置を手で与えることで、平行移動分の重なり領域を取得し、そこから対応点を求めることで対応点を安定的に取得する。本研究では画像を計測する工程が最も手間のかかるボトルネックとなるため、重なり領域を大きくするよう計測回数を増やすより、平行移動量を人手で設定する方が手間が小さいといえる。そして、この場合画像の結合処理において十分ずれが小さいと考えられる。

よって、本システムでは特にブレンディングに着目し開発を行う。4.1.3 で述べた手法において、重要となるのはブレンディング時の境界である。PIE においては境界が閉じている必要があり、本研究で対象としている異なる勾配をもつ二つの開領域を重ね合わせる処理には利用できない。そこで PB を利用し、その際にグラフ最小切断アルゴリズムを用いた境界の設定をすることで、高周波レイヤのずれを小さくし結合を行う。

4.3 PB における画像ピラミッドの取得

画像ピラミッドとは異なる解像度をもつ同一の画像集合であり、PB では画像の低周波成分を表す Gaussian ピラミッドと、エッジなどの高周波成分を表す Laplacian ピラミッドにより 1 枚の画像を表現しブレンディングを行う。

4.3.1 Gaussian ピラミッド

低周波成分の画像ピラミッド G_0, G_1, \dots, G_N は画像に対し REDUCE 処理と定義される重みカーネルの畳み込みとサブサンプリングを繰り返すことで得られる。

G_0 は元画像であり、 G_{i-1} から G_i が順次求められていく。図 4.4 は 1 次元での REDUCE 処理の例であり、 a, b, c はそれぞれカーネルの重みである。サイズ 5 のカーネルであるため、 $2^{N+1} \times 2^{N+1}$ サ

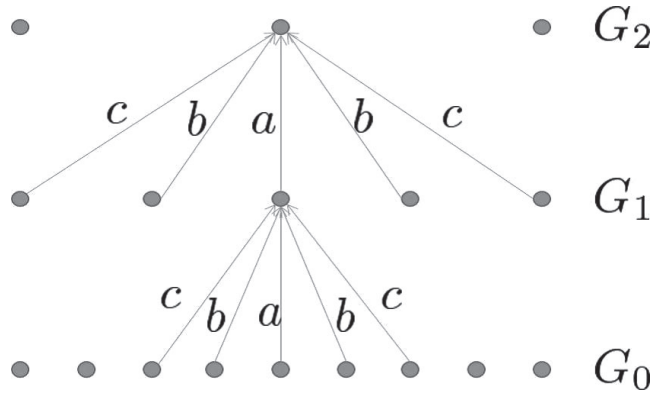


図 4.4 1次元における REDUCE 処理の例

イズの画像であれば $N+1$ のレイヤが生成される。

REDUCE 処理は次式により定義される。

$$G_l = REDUCE[G_{l-1}]$$

$$G_l(i, j) = \sum_{m,n=1}^5 w(m, n) G_{l-1}(2i + m, 2j + n) \quad (4.1)$$

ただし (i, j) は画像中の対象画素であり、 w は重みカーネルである。また、重みカーネルの設定として次式の制約条件が付く。

$$w(m, n) = \hat{w}(m) \hat{w}(n) \quad (4.2)$$

$$\hat{w}(0) = a, \hat{w}(\pm 1) = b, \hat{w}(\pm 2) = c \quad (4.3)$$

$$a + 2b + 2c = 1 \quad (4.4)$$

$$a + 2c = 2b \quad (4.5)$$

式 4.2 は次元に対し分割可能であること、式 4.3 はカーネルの対称性、式 4.4 は正規化を表している。式 4.5 は l レイヤの各重みの総和が $l+1$ の各ノードの重みと一致することから求められる。

よって a を任意の値としたとき $b = 1/4, c = 1/4 - a/2$ となる。 $a = 0.4$ の場合、カーネルは Gauss の確率密度関数となる。画像処理において Gaussian フィルタはローパスフィルタの役割を果たし、画像の低周波成分を抽出することになる。

そのため低周波成分の画像ピラミッドは Gauss 関数を重みカーネルとして用いた REDUCE 処理により得られ、これを Gaussian ピラミッドと呼ぶ。Gaussian Pyramid は各レイヤの REDUCE 処理にローパスフィルタの働きがあるため、全体としてバンドパスフィルタの役割をもつ。

4.3.2 Laplacian ピラミッド

Gaussian Pyramid の G_{l-1} と G_l の差分は G_l で通過しなかった成分を表している。そこで、各レイヤにおいて EXPAND 処理を行い、その差分から画像ピラミッドを作成する。 $k > 0$ に対し EX-

PAND 処理は以下で定義される。

$$\begin{aligned} G_{l,k} &= EXPAND[G_{l,k-1}] \\ G_{l,k}(i,j) &= 4 \sum_{n,m=-2}^2 G_{l,k-1}\left(\frac{2j+m}{2}, \frac{2j+n}{2}\right) \end{aligned} \quad (4.6)$$

ただし $G_{l,0} = G_l$ である。この時 $0 < l < N$ において L_0, L_1, \dots, L_N を次式により計算する。

$$L_l = G_l - EXPAND[G_{l+1}] = G_l - G_{l+1,l} \quad (4.7)$$

L_l は画像処理における高周波成分であるエッジ抽出に用いられる Laplacian フィルタと同様の作用をもち、 L_0, L_1, \dots, L_N を Laplacian ピラミッドと呼ぶ。

G の差分により L を計算しているため二つのピラミッドは次式の特性をもつ。

$$G_0 = \sum_{l=0}^N L_{l,l}. \quad (4.8)$$

レイヤごとに処理を行っても元画像である G_0 が復元可能となる。図 4.5 は $N=4$ における時の各ピラミッドの関係を図示したものである。ただし L は見やすさのため画素値を一様に加えている。ブレンディングで利用する際には、周波数の異なる各レイヤで高周波ほど急峻にブレンドを行い画像を統合することができる。

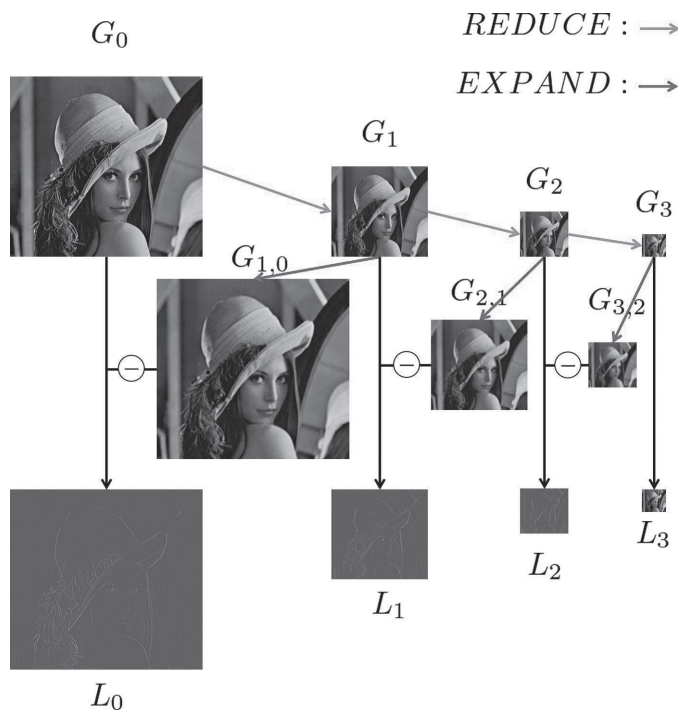


図 4.5 各画像ピラミッドの関係

4.4 グラフカット [22]

グラフカット法はグラフの最小切断アルゴリズムを使ったエネルギー最小化の手法である。画像をグラフとみなし、その最小切断を計算することで、領域分割や物体抽出、画像合成などの分野で広く利用されている。

画像の各画素をノード、画素間をエッジとしてグラフにし、そのエネルギー式を最小にするような切断を行う。エネルギー式は以下のように表わされる。

$$\begin{aligned} E(f) &= E_{data}(f) + E_{smooth}(f) \\ &= \sum_{p \in P} D_p(f_p) + \sum_{p, q \in N} V_{p, q}(f_p, f_q) \end{aligned} \quad (4.9)$$

ここで、 P はピクセルなどのサイトを表す集合、 $N \subset P \times P$ はサイト間の隣接関係を表す集合である。そして、 P の各サイトに有限個のラベルを割り振る。ラベルの集合を L として、 f は P の各サイト p にラベル $l \in L$ を与える写像である。また、式 4.9 のようなエネルギーはマルコフ確率場 (MRF) の最大事後確率推定の際によく見られる式である。 $E_{data}(f)$ と $E_{smooth}(f)$ はそれぞれデータ項と平滑化項と呼ばれ、データ項はベイズの式の尤度 $P(D|X)$ 、平滑化項はベイズの式の事前確率 $P(X)$ に対応する。

グラフカット法では最大流問題として式 4.9 のエネルギー式を最小化する。ここで、最大流問題と最小切断問題は線形計画法における双対にあたり、最大流最小切断定理によって最大流と最小切断は一致するとされている。

最大流問題を解くアルゴリズムに関して、Boykov [21] らによってライブラリが作成、公開されている。最大流問題を解く原型としては 2 種類に分けられ、1 つは Ford と Fulkerson の augmenting path アルゴリズムで、もう 1 つは Goldberg と Tarjan の push-relabel アルゴリズムである。Boykov らは augmenting path アルゴリズムを改良し多項式時間で解くアルゴリズムを提案している。

グラフカットの画像処理の応用例として、ノイズ除去や領域抽出、テクスチャ合成などがある [19]。文献 [20] では画像合成時の 2 画像間の境界線のずれを最小化するために合成先画像と新規画像の分割する境界をまたぐ 2 画素の距離の差を平滑化項に設定することで、境界の左右で最も画素値の差が小さい分割を行っている。

そのため本研究においてもこの手法が有用であると考え今回利用した。

4.4.1 グラフカットを用いた 2 画像間の境界設定

図 4.6 の様に既出画像を A 、新規画像を B とした時に、重なり領域で切断を行い、その cut から A 側であればそのまま、 B 側であれば新規画像を上書きする。これをグラフ表現に直すと、図 4.6 下のようになる。ある cut をまたぐノード間で A 側に属しているノードの位置を s 、 B 側に属しているノードの位置を t とすると、平滑化項は次式のようになる。

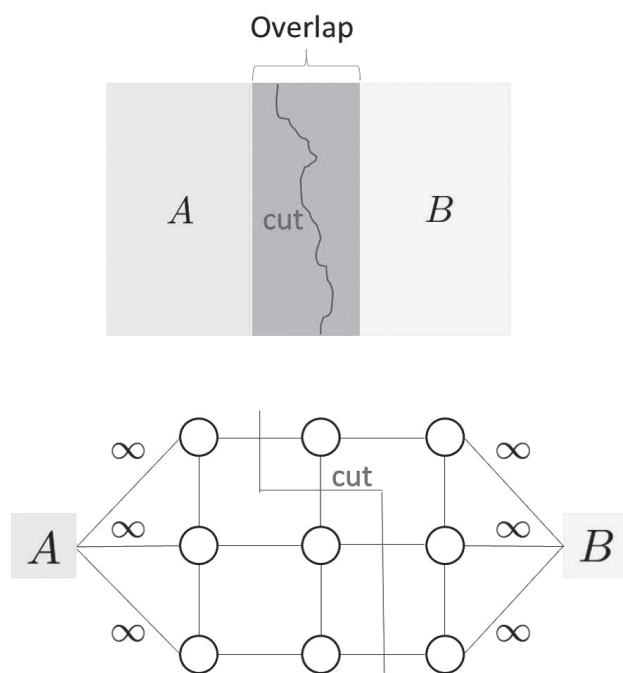


図 4.6 画像の重なりとグラフのカット

$$M(s, t, A, B) = \|A(s) - B(s)\| + \|A(t) - B(t)\| \quad (4.10)$$

$\|\cdot\|$ は画素間の 2 乗和としている。 M が平滑化項であり、 s, t 間のエッジを切断するコストとなる。これを最小化する cut を選ぶことで、境界のまた重なり領域内での切断となるよう A, B それぞれに接続しているノードにはデータ項として大きな値を設定し切断が起きないようにしている。それ以外ノードにはデータ項に値を与えず重なり領域内の任意の切断を行える。

⑤……………実装

本章では作成したシステムについて説明する。まず Kinect- 大型ディスプレイによる展示システムの実装、そして全形のサンプル画像群生成システムの実装について述べる。

5.1 Kinect と大型ディスプレイを用いたインタラクティブ展示

今回、全形におけるインタラクティブ展示システムの実装として大型ディスプレイと Kinect を用いた(図 5.1)。ディスプレイは 50 インチのフル HD の画面解像度である。Kinect のヘッドトラッキングによりディスプレイに対する観察者の頭の位置を取得し、ディスプレイの中心を原点とし頭の方角に応じた線形補間レンダリングを行う。その際、光の鏡面反射を考慮し頭の方角ではなく頭の方角と光軸に対象な方向の光源の出力を与える。また、システム自体は Kinect for windows sdk を用いて WPF として作成した。

3.4.2 でも述べたように 700mm × 400mm の領域であれば良好な出力が得られているためそれよりサイズの小さいディスプレイであればそのまま 1 度の計測で展示に利用できると考えられる。し



図 5.1 Kinect と大型ディスプレイによるインタラクティブ展示例

かし、4k ディスプレイ等のより dpi の高いディスプレイでは、より小さな領域で計測回数を増やす必要があるといえる。

5.2 全形サンプル画像群生成システムの構成

ここでは全体の処理をサンプル画像群抽出フェーズと全形サンプル画像群合成フェーズに分けて説明を行う。

5.2.1 部分サンプル画像群抽出フェーズ

このフェーズでは計測動画から光源方向に紐づくサンプル画像群の抽出までを行う（図 5.2）。

(a) 計測

計測時の条件としては以下の 3 点のようになる。

1. 撮影中手持ち光源以外に強い光源が存在していない（反射光を含む）
2. 撮影中手持ち光源を極端に試料に近づけすぎず、もしくは離し過ぎない
3. カメラが平面の試料に対しほぼ垂直である

厳密にこれらの条件を満たすことはないが、実験において得られるサンプル画像群に大きく影響が出たことはない。しかし、条件 2 についてはなるべく試料に対し一定の距離から照射することが望ましい。

(b) 初期化

サンプリングを行いたい光源方向を設定しマーカの位置を人手により設定する。マーカの設定については動画フレームからマーカ付近を拡大しマウスとキーボードにより 1 ピクセル単位で設定を行うインタフェースを用意した。

(c) フレーム抽出

3 章の手法を用いて設定された光源方向と最も近い動画フレームのインデックスを取得し、最終的に抽出を行う。その際、鏡面球のハイライトの最大輝度値が閾値以下の場合にはそのフレームを除外している。これはカメラの前を通過もしくは光源を一度試料から外す場合などを考慮した処理であり、閾値は実験的に 150 としている。フレーム抽出時にはマーカの内側部分を切り出して出力

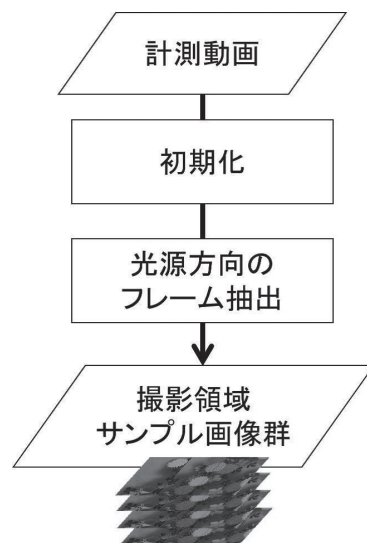


図 5.2 部分サンプル画像群抽出フェーズの流れ

を行う。以上の流れで、計測領域に対するサンプル画像群の抽出フェーズが終了する。そして、このフェーズを試料の大きさに合わせ繰り返し行い、5.2.2 のフェーズに移る。

5.2.2 全形サンプル画像群合成フェーズ

このフェーズでは複数の部分サンプル画像群から全形サンプル画像群合成を行う（図 5.3）。

(a) 初期化

ファイルの準備など諸設定を行う。画像の合成時は部分サンプル画像群の入ったフォルダの通し番号により行っているため合成の順番などを考慮しフォルダを構成する。最初の画像の平面を基準として変換行列の計算を行う。

(b) 平行移動量設定

図 5.4 のインタフェースにより画像をマウスにより並べる。

(c) 対応点・変換行列の取得

平行移動量から重なり領域を抽出し対応点は KLT 法を用いて対応点の取得を行う。

変換行列は各光源方向において共通で使用するため、例えば光源が試料垂直方向時の部分画像で対応点を取得し変換行列を求めれば、他の光源方向については同じ変換行列で結合を行える。

変換方法は平面射影変換を用いた。変換行列の計算には画像処理ライブラリである OpenCV の関数を利用し、RANSAC アルゴリズムにより多少のアウトライヤを含んだ対応点でもロバストな計算を行える。

(d) 最小切断の計算

既出画像に合うよう新規画像を変換した後に重なる領域を矩形として抽出し、その領域内の最小切断を求める。グラフの最小切断計算は Boykov らが公開しているライブラリを用いて計算を行った。

その際文献 [20] では、矩形の重なりを対象としている、本研究では変形後の画像を用いている

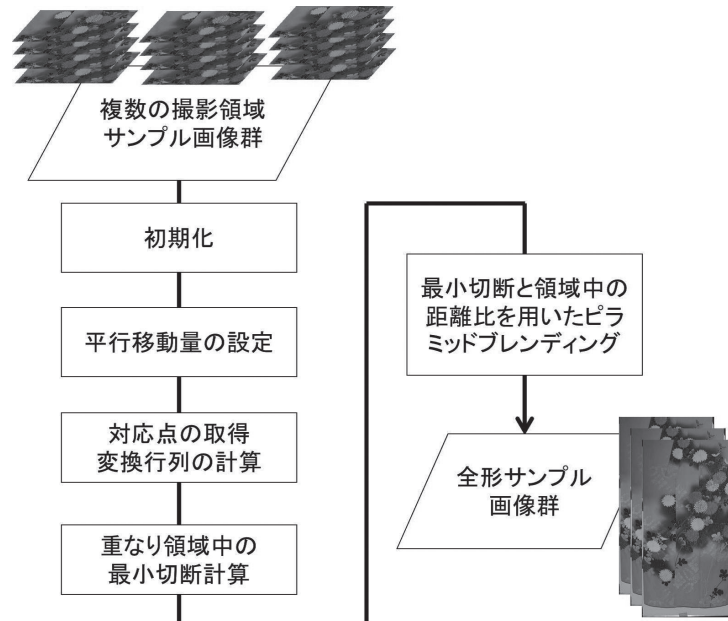


図 5.3 全形サンプル画像群合成フェーズの流れ

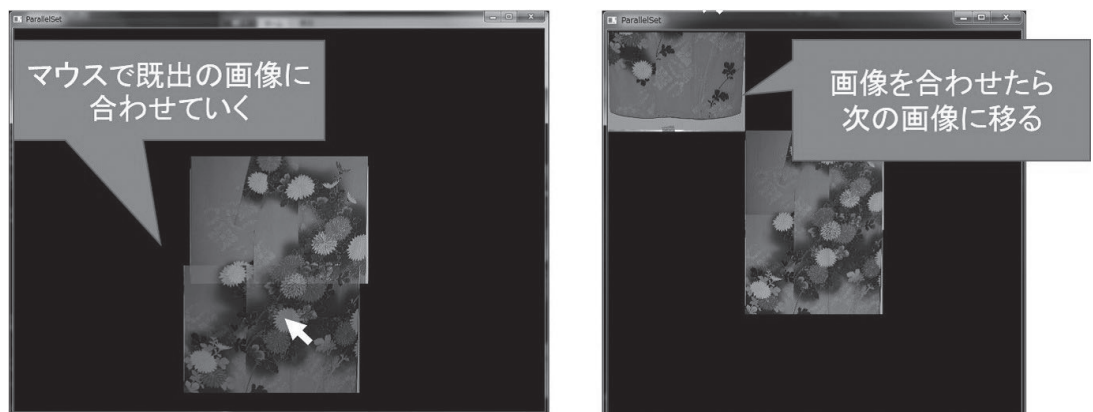


図 5.4 平行移動量の設定用インタフェース

ため図 5.5 の赤枠のような形になる。ライブラリの都合により、実装上では黒い枠を抽出し重なっていない、または画素値が存在しない領域中の画素間には平滑化項に大きな値を与え、赤い領域内で切断が起きるよう調整を行っている。

(e) ピラミッドのブレンディング

画像ピラミッドはカーネルサイズ 5 の Gaussian フィルタを用いて作成しレイヤ数は 5 としている。ブレンディング時の境界として L_0 は最小切断による境界を利用し、 L_1, L_2 については重なり領域において重なっていない最近接画素の距離の短さの比から α ブレンドを行っている。つまり図 5.6 の点 P に対し $d_B/(d_A+d_B)$ を既出画像 A の $L_1 \sim L_5$ の α 値、 $d_A/(d_A+d_B)$ を新規画像 B の $L_1 \sim L_5$ の α 値としている。

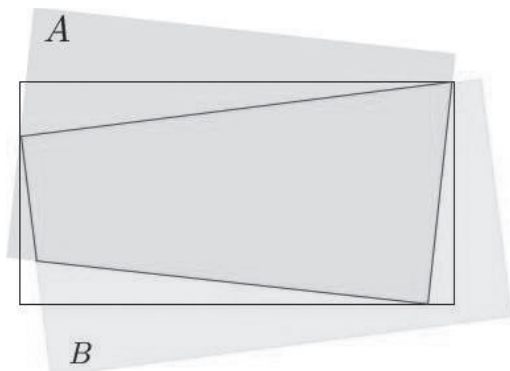


図 5.5 重なり領域に外接する矩形

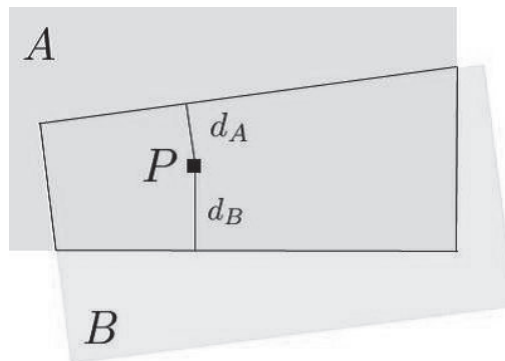


図 5.6 重なり領域内の点 P と画像 A, B との最近接距離

⑥……………実験

実験では、イメージモザイクにより合成したサンプル画像群の確認と、それを用いた大型ディスプレイの出力確認を行った。

6.1 実験 1：各種ブレンディング手法における結果比較

6.1.1 実験設定

今回は実験として図 6.1 の着物を対象に部分領域のサンプル画像群の取得からイメージモザイク処理までを行った。撮影条件や用いたデバイスは 3.4.2 の通りである。本来は全領域を計測する予定であったが、まずディスプレイに対して十分なサイズとして 3 つの部分領域において撮影を行い、全形サンプル画像群として確認を行った（図 6.1 右）。

結果の比較として、ブレンディングを行わずそのまま重ねた場合、距離の比による α 、ブレンディングせず最小カットで合成した例、そしてピラミッドブレンディング併用した本システムで出力を行った。

また実験時の詳細な各種装置・パラメータ設定を表 6.1 にまとめる。

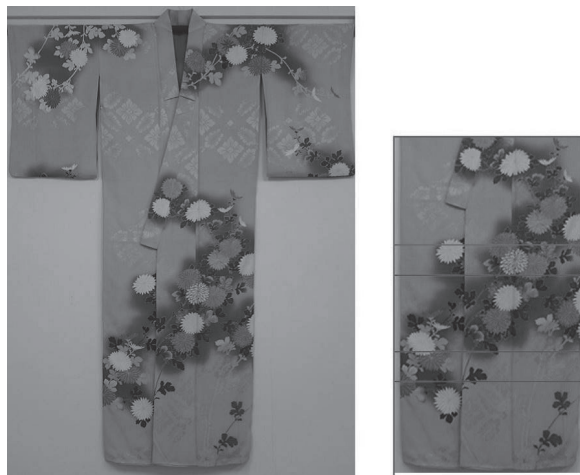


図 6.1 実験用試料として利用した着物

表 6.1 実験時のシステムの詳細な設定

カメラ	CanonEos6D
露出	SS: 50 絞り: 4.5 ISO: 1,600
レンズ	焦点距離: 100mm
ライト	色温度: 約 5,600 [K] 中心照度: 約 2,100 [Lx/50cm] 照射度: 約 40 度
マーカ	半径 20mm のプラスチック製半球
カメラと試料の距離	2,150mm
撮影領域サイズ	700 × 400mm
動画サイズ	1,920 × 1,080 (フル HD)
光源方向のサンプル間隔	θ 30 度, θ 45 度 (計 25 サンプル)
マーカ内の最低ハイライト輝度の閾値	150
ピラミッドブレンディングのレイヤ数	5

イメージモザイク処理においては合成画像が自然である・きれいであるといった評価を定量的に行うことは難しい。これはイメージモザイクにより得られる最終的な出力が重要であることから、処理中に発生する各画像の変形量や輝度補正值、または重なり合う領域内の画素値の距離や境界上のずれなどを定量的に求めてもそれが人間の感覚に一致しない場合が多いためである。

そのため目的に応じた合成結果が出力されればよいブラックボックス的な側面がある。さらに本研究ではインタラクティブ展示へ利用するため、合成した画像が、展示にどのような影響を及ぼすかを定量的に評価することは難しいため、ここではゴーストの発生やエッジのずれ、滑らかさなどを定性的に評価するにとどめる。

6.1.2 結果・考察

それぞれ出力結果を図 6.2, 図 6.3, 図 6.4 に示す。

図 6.2 から形状としてのずれはそもそも非常に小さい。平行移動量を設定せずに各画像間に対応点を計算した場合、対応点が正しく求まらず変換が行えなかったため、平行移動量を手動で設定し対応領域を限定することで安定した変換行列の取得が行えることが分かった。

図 6.3 を見ると、(a) において赤枠の部分にずれが発生しているが、(b) では別の境界をたどることでずれの発生を回避しておりエッジ上のずれはほとんど見られなかった。そのため、平面射影変換により残る小さなずれであればグラフカット法を用いることで解消できることが分かった。

図 6.4 では、赤枠の部分にゴーストの発生がみられた。(c) に比べ (d) ではゴーストの発生を大きく削減できていることが分かった。距離の比による α ブレンディングと、距離の比によるピラミッドブレンディングはほぼ同等である。そのためピラミッドブレンディング時に高周波レイヤの境界をグラフカットで設定することでゴースト発生を抑えつつ境界のずれの小さい画像が得られることが分かった。

一方で、ゴーストを完全に除去できてはいない。ピラミッドブレンディングではレイヤ数によりゴーストの発生するレイヤが異なり各レイヤのブレンディング処理を切り替える基準などが明確でないため適切なパラメータ設定が難しいと考えられる。そのため、今後パラメータを変えての再実験と理論的な解析を詳しく行っていきたい。

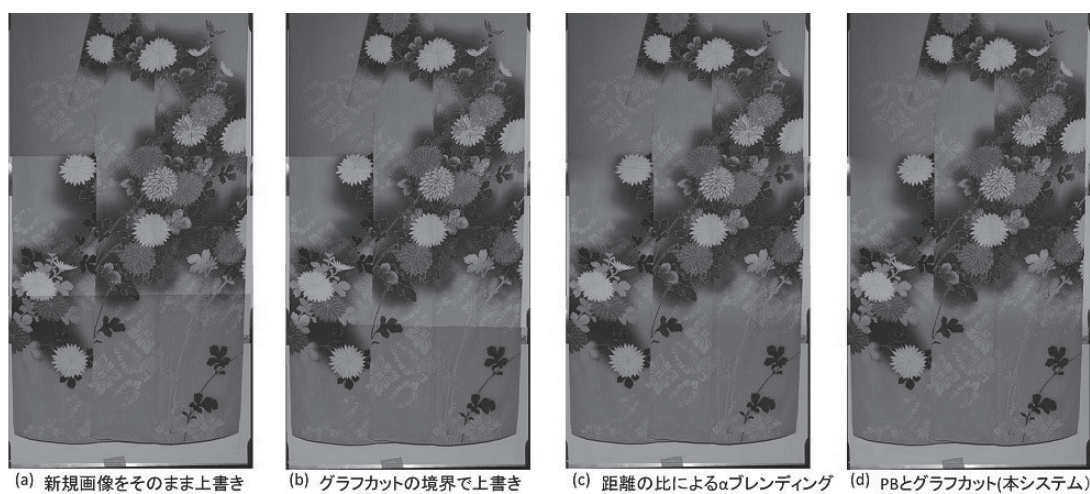


図 6.2 合成結果

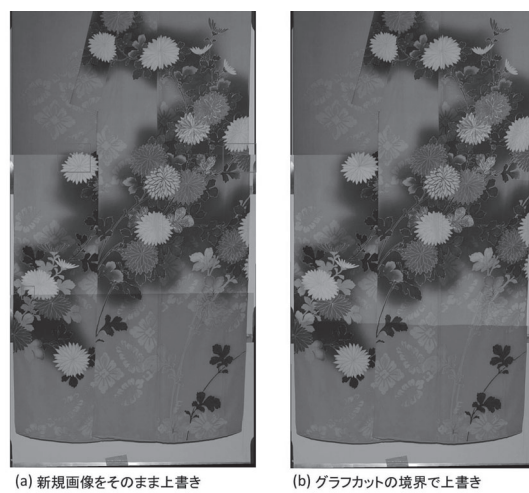


図 6.3 結果比較 1

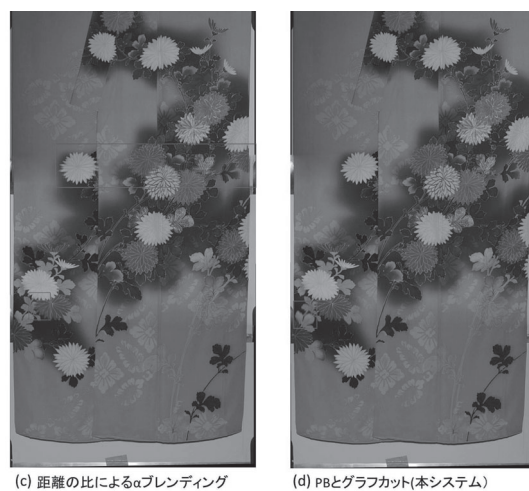


図 6.4 結果比較 2

7……………結論

本研究では小袖屏風を対象とし大型ディスプレイにおける全形のインタラクティブ展示システムの開発を行った。その際課題となる全形サンプル画像群の取得において簡便な方法の検討を行った。

そこで本研究では、手持ち光源により自由に光源を当て撮影した動画から、鏡面球マーカによるサンプル光源方向に紐づく画像群の抽出を行い、全形を復元するため試料の部分領域ごと得られたサンプル画像群をイメージモザイクにより結合した。実験ではサンプルする光源間隔に対して誤差率約8%の精度で光源方向を取得することが可能になった。またグラフカットとピラミッドブレンディングを用いたイメージモザイク処理によりずれやゴーストを抑えた合成サンプル画像群を取得し、大型ディスプレイによるインタラクティブ展示が可能となった。これらの手法は、簡易な手段で資料の質感や存在感を表現できる可能性があり、小袖屏風以外の資料、例えば大判の地図や絵画等への応用も期待できる。

このようにインタラクティブ展示は従来にない感性の再現が期待できるものの、そのリアリティについての定量評価の手法は明らかではない。今後は、引き続き実験により全計測によりサンプル画像群の取得を行い、部分領域ごとの明るさの違いを考慮したブレンディング手法の検討をしていきたい。また、システムのパラメータ設定をまとめ人の感覚に対する影響の定量的な評価を行っていきたい。

参考文献

- [1] 金谷一郎：“文化財画像計測”，計測と制御，Vol.47，No.1，pp.36-40（2008）。
- [2] 佐藤智和，横矢直和：“VR/MRのための画像計測”，計測と制御，Vol.47，No.1，pp.30-35（2008）。
- [3] 西脇靖洋，坂口嘉之，田中弘美：“動的照明による織物の異方性反射レンダリング”，電子情報通信学会技術報告，Vol.110，No.381（PRMU2010-149-207），pp.223-228（2011）。
- [4] 武田祐樹，田中弘美：“多方向照明：HDR画像を用いた金襴の多重解像度異方性BTFモデリング”，電子情報通信学会論文誌D情報・システム，Vol.J91-D，No.12，pp.2729-2738（2008）。
- [5] 佐藤いまり：“リアリティの再現”，情報処理，Vol.47，No.4，pp.346-352（2006）。
- [6] 小澤耕太郎：“高度歴史資料展示における高リアリティインタフェースの構築”，オープンソースモジュールスタジオ報告書（2013）。
- [7] 比留間伸行，中口俊哉：“心理物理モデルに基づく像評価”，電子情報通信学会誌，Vol.96，No.4，pp.228-233（2013）。
- [8] Y. Sakano and H. Ando: “Effects of head motion and stereo viewing on perceived glossiness”, Journal of Vision, Vol.10, No.9-15, pp.1-14（2010）。
- [9] Mirko Sattler, Ralf Sarlette and Reinhard Klein: “Efficient and Realistic Visualization of Cloth”, Proceedings of Eurographics Symposium on Rendering, pp.167-177（2003）。
- [10] Peter-Pike Sloan, Jan Kautz and John Snyder: “Precomputed radiance transfer for real-time rendering in dynamic, low-frequency lighting environments”, Proceedings of SIGGRAPH, Vol.21, No.3, pp.527-536（2002）。
- [11] 佐藤いまり，林田守広，甲斐郁代，佐藤洋一，池内克史：“実光源環境下での画像生成：基礎画像の線形和による高速レンダリング手法”，電子情報通信学会論文誌，D-II，情報・システム，II-パターン処理，Vol.J84-D-II，No.8，pp.1864-1872（2001）。
- [12] Y. Sakano and H. Ando: “Estimation of Illuminant Direction and Intensity of Multiple Light Sources”，

-
- Computer Vision - ECCV 2002, Vol.10, No.9-15, pp.206-220 (2002).
- [13] Bruce D. Lucas, Takeo Kanade: "An Iterative Image Resistration Technique with an Application to Stereo Vision", International Joint Conference on Artificial Intelligence, pp.674-679 (1981).
- [14] Carlo Tomasi, Takeo Kanade: "Detection and Tracking of Point Features", Carnegie Mellon University Technical report CMU-CS-91-132 (1991).
- [15] Jianbo Shi, Carlo Tomasi: "Good Feature to Track", IEEE Conference on Computer Vision and Pattern Recognition, pp.593-600 (1994).
- [16] 千葉直樹, 蚊野浩, 美濃導彦, 安田昌司: "画像特徴に基づくイメージモザイク", 電子情報通信学会論文誌 D-II, 情報・システム, II-パターン処理, J82-D-II (10), pp.1581-1589 (1999).
- [17] Peter J. Burt, Edward H. Adelson: "A multiresolution spline with application to image mosaics", ACM Transactions on Graphics Volume 2, Issue 4, pp.217-236 (1983).
- [18] PEREZ. P, GANGNET. M, AND BLAKE. A: "Poisson image editing", ACM Transactions on Graphics Volume 22, Issue 3, pp.313-318 (2003).
- [19] Alexei A. Efros, H. Ando, William T. Freeman: "Image quilting for texture synthesis and transfer", Proceedings of the SIGGRAPH, ISBN1-58113292-1, pp.341-346 (2001).
- [20] Vivek Kwatra, Arno Schodl, Greg Turk, Aaron Bobick: "Graphcut textures: image and video synthesis using graph cuts", ACM Transactions on Graphics Volume 22, Issue 3, pp.277-286 (2003).
- [21] Y. Boykov and V. Kolmogorov: "An Experimental Comparison of MinCut/Max-Flow Algorithms for Energy Minimization in Vision", IEEE Trans. Pattern Analysis and Machine Intelligence 26 (9), pp.1124-1137 (2004).
- [22] 八木康史, 斎藤英雄, 石川博: "コンピュータビジョン最先端ガイド1", アドコム・メディア株式会社, pp.39-74, 初版 (2008).

(横浜国立大学大学院工学研究院, 国立歴史民俗博物館共同研究員)

(2019年3月14日受付, 2019年8月5日審査終了)

Development of an Interactive Exhibition System for Kosode Byobu

HAMAGAMI Tomoki

This study develops a new interactive exhibition system for Kosode byobu by using a simple re-active effects display method. For realistic experience for exhibition, the reproduction of space with objects and appreciators is a significant issue. Especially it is known that local illuminations such as BRDF, BSSRDF, and BTF methods are effective technique. However, the data acquisition cost from valuable historical objects becomes vast large.

This study proposes a simple image collection method by using a hand-held camera and a spherical marker, then constructing reproduction images with image mosaic processing by using pyramid building and graph-cut. The experiment results show the proposed method enables the reproduce image within an 8% error, and realtime rendering in a large display.

Key words: interactive, pyramid building, ray tracing, BRDF, BTF