

ウェアラブル AR のための 注釈と対象物体の視認性を考慮したビューマネジメント

牧田 孝嗣 神原 誠之 横矢 直和
奈良先端科学技術大学院大学 情報科学研究科

あらまし 拡張現実感を利用して現実環境中に存在する物体に関する注釈を提示する場合、注釈の配置方法の工夫（ビューマネジメント）を利用することで、注釈の理解のしやすさを向上可能である。ビューマネジメントを用いた注釈付けを実現するには、注釈を付加する対象物体の実環境中での位置や形状が必要となるが、対象物体が移動体や非剛体である場合、それらを正確に得ることは困難である。そこで本研究では、シーンに存在する移動体や非剛体を注釈対象として、その存在領域推定に基づいて注釈と対象物体の視認性を考慮した注釈のビューマネジメント手法を提案する。提案手法では、ネットワークを介して獲得した位置と対象の形状を近似したプリミティブを用いて対象物体の存在領域の推定を行う。さらに、注釈と対象物体の視認性を下げる要素である注釈と対象物体などの重なり、注釈と対象物体との距離、注釈のフレーム間の移動量をペナルティとして定量的に定義し、これらを最小化することでビューマネジメントを行う。プロトタイプシステムを用いて、環境中の複数のウェアラブルコンピュータのユーザに対する注釈付け実験を行い、提案手法の有用性を示す。

View Management Based on Visibility of Annotation and Target Object for Wearable Augmented Reality

Koji MAKITA, Masayuki KANBARA, and Naokazu YOKOYA
Graduate School of Information Science, Nara Institute of Science and Technology

Abstract: This paper proposes a new view management method based on visibility of annotation and target object for wearable augmented reality (AR). In annotation overlay applications using AR, a view management is widely used to improve readability and intelligibility of annotations in AR image. To recognize visible portions of objects on the user's view, a position, an orientation and shape of objects are known in conventional view management methods. However, it is difficult for the wearable AR system to obtain positions, orientations and shapes of objects because there is usually a moving or nonrigid object. This paper proposes a view management method to overlay annotations of moving or nonrigid objects for wearable AR. In the proposed method, the wearable AR system obtains positions and shapes of target objects via a network for estimating the visible portions of target objects in the user's view. Annotations are located by minimizing penalties; an overlap of annotation, an occlusion of target objects, a distance between the annotation and the target object, and a distance of the annotation in the sequential frames. In experiments, the prototype system has been proven to correctly provide each user with annotations on multiple users of wearable systems.

1 はじめに

拡張現実感（Augmented Reality; AR）は、計算機が生成した情報を現実環境に重ね合わせて表示する技術である。装着して使用可能なウェアラブルコンピュータと、AR を組み合わせたウェアラブル AR を利用すれば、ユーザは任意の場所において、自己位置・姿勢に応じた位置依存情報を直感的に得ることができる。ウェアラブル AR を応用したシステムの例として、位

置依存情報を注釈として提示するものが挙げられる[1][2]。注釈を描画する場合には、現実空間中に存在する 3 次元モデルとして描画する方法や、ユーザ視界中に 2 次元画像を描画する方法がある。図 1 に、それぞれの方法を用いて作成された注釈付加画像の例を示す。注釈情報を提示する場合には、注釈の配置方法の工夫（ビューマネジメント）を用いることで注釈の見やすさを向上できるため、これまでに建造物や機器

のスイッチ・ボタンなどの物体に関する注釈情報を対象としたビューマネジメント手法が提案されている[3-6]。これらの手法はいずれも静的環境を前提とし、注釈の対象である物体の位置・姿勢・形状があらかじめ既知のデータとして与えられていた。そのため、注釈対象が移動体や非剛体が想定されるウェアラブルARのように、正確な対象の位置・姿勢・形状を得ることが困難である場合に利用可能なビューマネジメント手法が望まれる。

そこで本研究では、ウェアラブルARに代表されるようなシーンに存在する移動体や非剛体を注釈対象として、注釈と対象物体の視認性の向上を目的とした注釈のビューマネジメント手法を提案する。提案手法では、ネットワークを介して獲得した対象の位置情報と形状プリミティブを用いて対象物体の存在領域の推定を行う。さらに、注釈と対象物体の視認性を下げる要素である注釈と対象物体などの重なり、注釈と対象物体との距離、フレーム間の注釈の移動量をペナルティとして定量的に定義し、これらを最小化することでビューマネジメントを行う。本稿では、手法[7]により赤外線カメラを利用して頭頂部の位置・姿勢の計測が可能なウェアラブルコンピュータのユーザを注釈対象として想定する。さらに、ネットワーク共有データベースとハイブリッドP2Pを利用した注釈提示[8]のフレームワークを利用して注釈対象の位置を獲得する。

以下、2章では注釈と対象物体の視認性を考慮したビューマネジメントについて、3章ではプロトタイプシステムを用いた注釈付け実験について説明する。最後に4章では本稿のまとめと今後の課題について述べる。

2 注釈対象の存在領域推定に基づくビューマネジメント

2.1 ビューマネジメントの基本方針

現実環境に注釈を重畳表示する場合、先行研究では、対象物体の存在領域に重なる位置に描画する方法[2]、対象物体の存在領域外に注釈を描画し、対象領域と注釈を線で結ぶ方法[1][4][6]、またはそれら2つの手法を併用する方法[3][5]が一般的である。このような描画方法を利用する際、注釈と対象物体の理解を妨げる要素としては、注釈同士の重なりや注釈と対象物体の重なりといった、注釈のレイアウトに関するものや、注釈の動きに関するものなどがある。本研究で

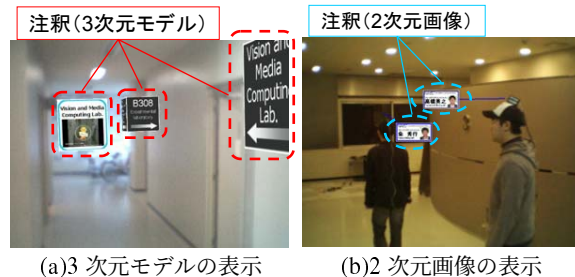


図1 注釈付加画像の例

は、対象物体と注釈を線で結ぶ方法を用い、以下の3つの観点をもとにしたペナルティを定義する。

- 注釈，線，対象物体に関連する重なり量（小さい方が低い）
- 対象物体と注釈を結ぶ線の長さ（短い方が低い）
- フレーム間における注釈の移動量（小さい方が低い）

これらの観点を考慮したペナルティを算出するためには、ウェアラブルARシステムのユーザの視界中に存在する注釈対象の存在領域を推定する必要がある。そこで、ビューマネジメントを用いた注釈付加を行う過程として、注釈対象の存在領域推定、ペナルティの算出、注釈の配置位置の決定の3段階を考える。

以下、2.2節では、提案手法における注釈対象の存在領域の推定方法、2.3節では、ペナルティの算出、2.4節では注釈の配置位置の決定について、それぞれ述べる。

2.2 注釈対象の存在領域推定

現実環境中に存在する注釈対象の存在領域を把握するために、提案手法では、注釈対象の現在位置と形状を利用した存在領域推定を行う。具体的には、図2に示すように、ユーザの視界において色の濃淡が注釈対象の存在確率を示す図を作成する。以降、この図を“存在確率マップ”と呼ぶ。ウェアラブルコンピュータのユーザは、手法[7]により位置・姿勢推定を行うための赤外線カメラを頭頂部に装備している。以下にウェアラブルARシステムのユーザが見る現実環境（以降、ユーザ視界）における注釈対象の存在領域の推定手順を述べる。

まず、注釈対象の現在位置を獲得するために、手法[8]のネットワーク共有データベースのフレームワークを利用する。ネットワークを介して移動オブジェクトの位置を獲得することで、ユーザ視界における注釈対象の位置を把握する。

次に、注釈対象の形状をボクセル（立方体）や直方

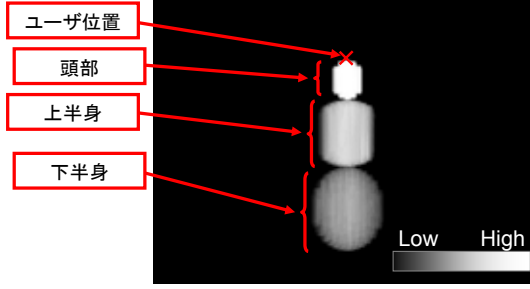


図2 存在確率マップの例

体などの立体形状の集合により構成されるプリミティブをネットワーク共有データベースより取得する。本稿では、ウェアラブルコンピュータのユーザは床面におおよそ鉛直に立っていることを想定する。まず、ユーザの体を大きく3つの部分（頭、上半身、下半身）に分け、次に各部分を立体形状の集合で表現する。さらに、各部分に存在確率を設定する。本稿では、ユーザ位置は頭頂部に装備されたカメラ位置として算出される。また、注釈の対象が人の場合には一般的に顔部分が隠蔽を回避する重要部分であると考えられることから、存在確率は頭頂部に近いほど大きい値を設定する。最後に、全ての立体形状をユーザ視界に投影する。その際、形状の色を存在確率に反比例した透明度（アルファ値）として表現することで、一般的なレンダリング処理により生成された画像が存在確率マップとして利用可能となる。

ただし、立体形状の集合は最終的に出力する注釈付加画像には提示しないため、立体形状の描画にはオフスクリーンレンダリングを用いる。

2.3 ペナルティの算出

画像上の位置 (x, y) に注釈を配置する場合のペナルティ $P(x, y)$ を、以下の式で定義する。

$$P(x, y) = w_o P_o(x, y) + w_l P_l(x, y) + w_m P_m(x, y) \quad (1)$$

$P(x, y)$ は、3つのペナルティ項目の和で定義される。上記の式における各ペナルティの内容は以下の通りである。

$P_o(x, y)$: 注釈と線を描画する際に発生する重なりに関するペナルティ

$P_l(x, y)$: 注釈と対象物体を結ぶ線の長さに関するペナルティ

$P_m(x, y)$: 前フレームにおける注釈の位置からの移動量に関するペナルティ

また、 w_o , w_l , w_m は各々のペナルティの重みの係数である。以下に、各々のペナルティ項目 P_o , P_l , P_m の算出方法についてそれぞれ述べる。

重なりに関するペナルティ P_o の算出：

重なりに関するペナルティは、注釈のレイアウトから算出されるペナルティと、注釈と対象物体の存在領域から算出されるペナルティの2種類のペナルティの和とする。以下に、各々のペナルティの算出の方法を述べる。

- 注釈のレイアウトから算出されるペナルティ

P_{layout}

注釈のレイアウトのみから算出されるペナルティは、(a) 注釈同士の重なり、(b) 線同士の重なり、(c) 注釈と線の重なり、の3つの重なり部分に含まれる画素の数の合計とする。図3に、各々の重なりが発生するような注釈の配置パターンの例を示す。

- 注釈のレイアウトと対象物体の存在領域から算出

されるペナルティ P_{map}

図4に、注釈のレイアウトと対象物体の存在領域から算出されるペナルティの発生する注釈の配置パターンの例を示す。対象物体の存在領域は存在確率マップによって推定されており、正確な存在範囲ではない。そこで、位置 (x, y) に注釈が配置された場合のペナルティ $P_{map}(x, y)$ を存在確率マップの持つ確率の和として定義する。具体的には以下の式で定義する。

$$P_{map}(x, y) = \sum_{i \in A_{(x, y)}} C_i \quad (2)$$

ただし、 i は画像上の点の座標、 C_i は存在確率マップ上の座標 i における画素の輝度値、 $A_{(x, y)}$ は、注釈が座標 (x, y) に表示された場合に隠蔽される画素の集合である。

線の長さに関するペナルティ $P_l(x, y)$ の算出：

注釈と対象物体を結ぶ線の長さに関するペナルティは、線の長さに比例して大きくなるような関数で定義する。本稿ではパラメータの設定を簡単化するために、 $P_l(x, y)$ を以下の式で設定する。

$$P_l(x, y) = l^2 \quad (3)$$

ただし、 l は注釈と対象物体を結ぶ線の長さである。

注釈の移動量に関するペナルティ ($P_m(x, y)$) の算出：

図5に、注釈の移動量の算出方法の概念図を示す。注釈が位置 $B(x, y)$ に配置される際の注釈の移動量に関するペナルティ $P_m(x, y)$ を以下の式で定義する。

$$P_m(x, y) = |B_t - A_t| \quad (4)$$

ただし、 A_{t-1} は1つ前のフレームにおける注釈の位置、 A_t は現フレームにおいて1つ前のフレームにお

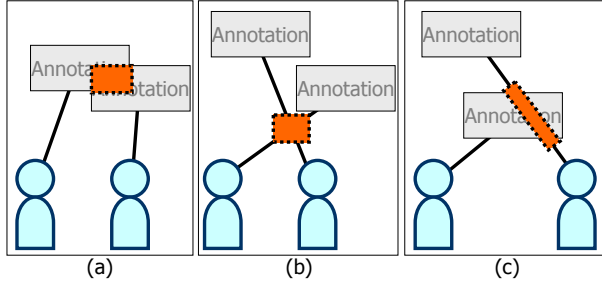


図3 ペナルティの発生する注釈の配置パターンの例

る注釈と対象物体の相対位置関係を保存して得られる位置、 $|B_l - A_l|$ は位置 B_l と位置 A_l との距離である。

2.4 ペナルティ最小化による注釈付加

2.3 節において定義したペナルティを用いてビューマネジメントを行い、注釈の付加を行う。注釈の配置パターンは、基本的には無数に存在し、各々のパターンについてペナルティを算出できる。そのため、ペナルティの最小化を行うために、以下の処理手順を用いる。

1. 注釈の大きさと形状を設定
2. 注釈を配置可能な領域中から、配置候補位置を設定
3. ペナルティが最小となる注釈の配置パターンを決定

以下、各手順について詳述する。

1. 注釈の大きさと形状の設定

様々な情報が含まれた注釈は、ウェアラブル拡張現実感システムに用いられる提示デバイスを利用してユーザに提示される。そのため、注釈の大きさや形状は、注釈の内容、及び提示デバイスの仕様を考慮した上で決定する必要がある。例として、文字列が表示された注釈を考える。この場合、提示デバイスの大きさを考慮した上で、ユーザにとって文字の読める注釈の大きさを確保する必要がある。一方、注釈を大きく設定することで、重なりに関するペナルティが大きくなる可能性が高くなる。以上の観点から、注釈の大きさ及び形状を適切に固定値で設定する。

2. 配置候補位置の設定

ビューマネジメントを行うために算出するペナルティは、注釈を配置する位置によって決まる。注釈の配置位置は基本的には無数に存在するため、本手法では計算の高速化を行うために各注釈を配置する位置の候補を複数設定し、その中からペナルティが最小となる位置に注釈を配置する。設定する位置の候補の数に制限はないが、ペナルティの計算に必要な時間は、候補数

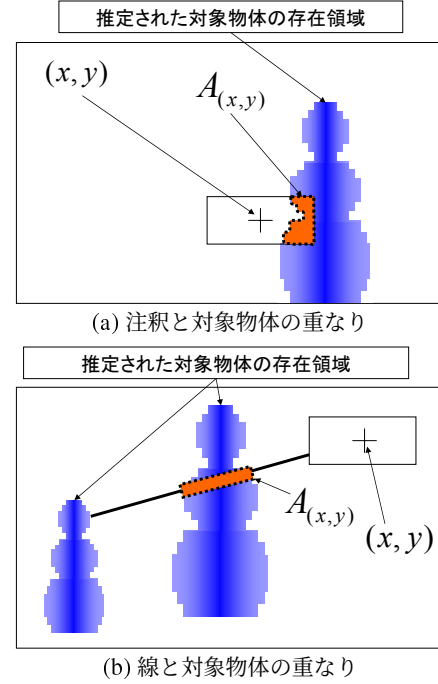


図4 対象物体との重なりによりペナルティの発生する注釈の配置パターンの例

に比例して大きくなる。

このような観点から、注釈を配置する候補の位置を、以下のように設定する。まず、注釈を配置する候補位置は、前フレームにおける注釈位置と対象物体の相対位置を保存した位置の周辺に限定する。ただし、候補位置を現在位置の周辺のみ限定すると、周辺にペナルティの小さい位置がなくなった場合に注釈がペナルティの大きい位置に描画され続ける場合がある。そこで、周辺領域に加えて、一定距離離れた位置を数点加え、配置位置の候補を設定する。図6に、一定距離離れた位置を8点加えた場合の、注釈を配置する候補位置の設定例を示す。

3. 注釈の配置パターンの決定

ペナルティの最小化には、全体最適化を用いる方法と、逐次処理による局所最適化を用いる方法がある。本研究では実時間での動作が要求されるため、逐次処理を行う。

次に、逐次処理による局所最適化を用いる方法について述べる。本手法では、各々の注釈の配置位置を、1つずつ順次決定する。この場合、ユーザ視界に付加する注釈の個数が n 個、各々の注釈の配置位置の候補数を p_i とすると、ペナルティの計算回数は $p_1 + p_2 + \dots + p_n = \sum_{k=1}^n p_k$ となり、全体最適化と比較して計算量は小さい。注釈の配置位置を逐次的に決

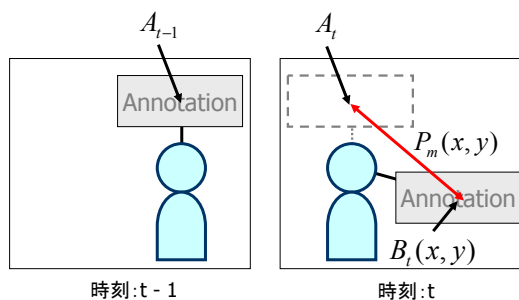


図 5 注釈の移動量の算出方法

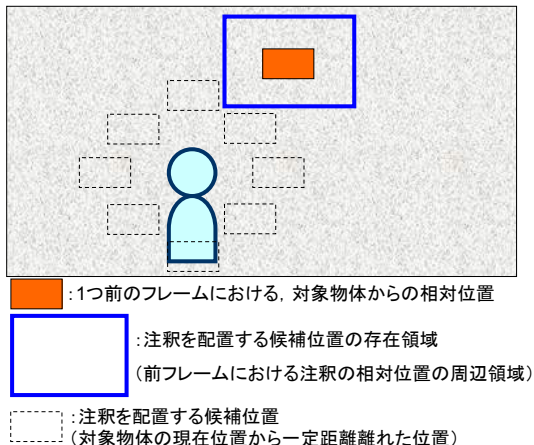


図 6 一定距離離れた位置を 8 点加えた場合
の、注釈を配置する候補位置の設定例

定するため、局所解によりペナルティが最小のパターンが出力されない可能性はあるが、ペナルティの大きさは最適解に近い値が得られると考えられる。

3 注釈提示実験

3.1 実験環境

本実験ではまず、現実環境中における位置・姿勢が固定された注釈がネットワーク共有データベースに登録されている環境を準備した。次に、手法 [7] により位置・姿勢の計測できる 4 人のウェアラブルシステムのユーザ (A, B, C, D) の位置を双方向に P2P で送信し、ユーザ A が装着したビデオシースルー型のウェアラブル注釈提示システムを用いて注釈対象の存在領域推定を行った。表 1 に、それぞれのシステムに使用した機器の構成を示す。今回使用した手法 [7] におけるユーザ A, B, C の位置誤差平均は約 9mm、姿勢誤差平均は約 0.35 度であり、ユーザ D の位置誤差平均は約 11mm である。また、ユーザ A, B, C の推定頻度は 1 秒あたり約 29 回、ユーザ D の推定頻度は 1 秒あたり約 30 回である。本実験において、各ユーザはいずれも頭頂部に位置・姿勢の推定用の赤外線カメラを装着しており、ユーザの存在領域は、水平面に鉛直な直方体の集合で推定している。ユーザの体の構造上、

表 1 各システムの機器構成

	使用機器	スペック
ウェアラブルシステム (ユーザ A)	計算機	CPU: 2.13GHz メモリ: 1.0GB 54Mbps Ethernet 接続
	位置・姿勢計測用 赤外線カメラ	オリジナル 有効画素数: 1024 × 768 画像更新速度: 約 29fps
	ユーザ視界画像 取得用カメラ	オリジナル 有効画素数: 640 × 480 画像更新頻度: 約 15fps
ウェアラブルシステム (ユーザ B, C)	計算機	CPU: Core2Duo, 2.0GHz メモリ: 3GB 54Mbps Ethernet 接続
	位置・姿勢計測用 赤外線カメラ	オリジナル 有効画素数: 1024 × 768 画像更新速度: 約 29fps
ウェアラブルシステム (ユーザ D)	計算機	CPU: 1.2GHz メモリ: 500MB 54Mbps Ethernet 接続
	位置・姿勢計測用 赤外線カメラ	オリジナル 有効画素数: 640 × 480 画像更新速度: 約 30fps

頭部はほぼ剛体であり、カメラ位置も近いことから、アルファ値を高い値で設定する。また、上半身及び下半身は手足の動きによって様々な写り方をするため、アルファ値を低い値で設定するものとする。本実験では、描画に必要な計算時間の短縮のために、ボクセル (立方体) を用いる代わりとして直方体の集合を用いる。ユーザの各部分の存在領域を推定する直方体の個数及びアルファ値 (透過係数) はそれぞれ、頭部 (個数: 25 個, アルファ値: 0.6), 上半身 (個数: 81 個, アルファ値: 0.1), 下半身 (個数: 169 個, アルファ値: 0.05) と設定している。

3.2 ビューマネジメントを用いた注釈付け

注釈対象の存在領域推定で得られた結果を利用して、注釈付加画像の作成実験を行った。本実験ではまず、640 画素 × 480 画素の注釈付加画像を作成することとし、作成した注釈用の画像の内容 (顔写真、所属及び名前を示す文字列) を考慮して、注釈画像を 81 画素 × 41 画素の長方形に設定した。また、ある位置 (x, y) に注釈を配置する場合のペナルティ $P(x, y)$ の重み付け係数を、それぞれ $w_o = 0.1$, $w_l = 0.5$, $w_m = 0.2$ と設定した。ただし、線同士の重なり、線と注釈の重なり、線と対象物体の重なりの量を計算するために、線からの距離が 3 画素以内の画素を、線上にある画素とした。

次に、注釈を配置する候補位置の設定を行った。本実験では、前フレームにおける注釈と対象物体の相対位置関係を保存して得られる位置を中心として、周辺の 21 画素 × 21 画素の範囲を、候補位置として設定した。さらに、対象物体の現在位置から $(450 \div l)$ 画素

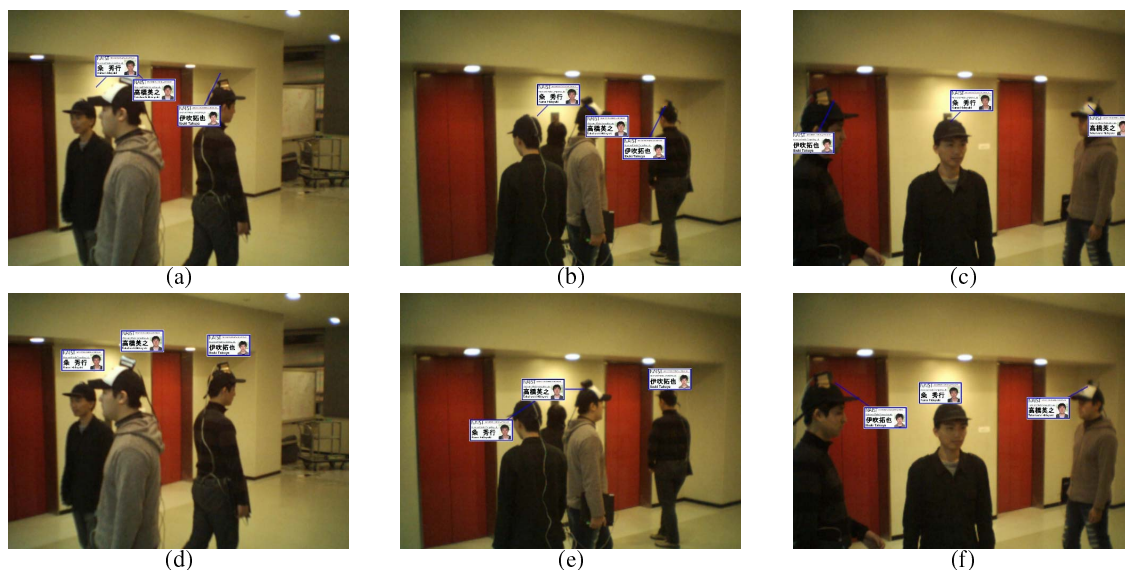


図7 ビューマネジメントの有無の比較結果

離れた円周上に16点の候補位置を設定した。ただし l は、現実空間中における対象物体の位置とウェアラブル拡張現実感システムの位置との距離である。

図7に、ビューマネジメントの有無の比較結果を示す。図7(a)(b)(c)はビューマネジメントを利用せず、注釈とユーザとの相対位置を固定して作成された注釈付加画像である。また、図7(d)(e)(f)は、提案手法を用いて作成された注釈付加画像である。ビューマネジメントを行わない場合には、注釈とユーザの重なりが確認できる。しかしビューマネジメントを行った場合には、ビューマネジメントにより注釈の位置を動的に変更しているため、注釈同士の重なり及び注釈とユーザの重なりを避けた注釈付けができています。なお、本実験では注釈の位置決定順序をあらかじめ固定した逐次処理を用いており、動作フレームレートはおよそ8～12フレーム毎秒であった。

4 まとめ

本研究では、ウェアラブルARを利用した注釈提示において、注釈と対象物体の視認性を考慮したビューマネジメント手法を提案した。また、ウェアラブルコンピュータのユーザを対象とした実験を行い、ペナルティ最小化を利用したビューマネジメントが可能であることを示した。今後は、3Dオブジェクトによる注釈との混在環境におけるビューマネジメント手法への拡張を行う予定である。

謝辞 この研究の一部は、特別教育研究経費「超高速ディペンダブルネットワーク環境での複合現実感研究創出事業」の支援による。

参考文献

- [1] J. Mooser, S. You and U. Neumann: "Real-Time Object Tracking for Augmented Reality Combining Graph Cuts and Optical Flow," sixth IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR 2007), pp. 145-152, 2007.
- [2] 天目 隆平, 神原 誠之, 横矢 直和: "ウェアラブル拡張現実感システムのための注目オブジェクトへの直感的な注釈提示手法," 日本バーチャルリアリティ学会論文誌, Vol. 10, No. 3, pp. 305-312, 2005.
- [3] B. Bell, S. Feiner and T. Höller: "View Management for Virtual and Augmented Reality," ACM Symp. on User Interface Software and Technology (UIST 2001), pp.101-110, 2001.
- [4] R. Azuma and C. Furmanski: "Evaluating Label Placement for Augmented Reality View Management," Proc. 2nd Int. Symp. on Mixed and Augmented Reality (ISMAR 2003), pp. 66-75, 2003.
- [5] F. Zhang and H. Sun: "Dynamic Labeling Management in Virtual and Augmented Environments," 9th International Conference on Computer Aided Design and Computer Graphics(CAD/CG 2005), pp. 397-402, 2005.
- [6] F. Shibata, H. Nakamoto, R. Sasaki, A. Kimura and H. Tamura: "A View Management Method for Mobile Mixed Reality Systems", Proc. 14th Eurographics Symp. on Virtual Environments (EGVE2008), pp. 17-24, 2008.
- [7] 中里 祐介, 神原 誠之, 横矢 直和: "ウェアラブル拡張現実感のための不可視マーカと赤外線カメラを用いた位置・姿勢推定," 日本バーチャルリアリティ学会論文誌, Vol. 10, No. 3, pp. 295-304, 2005.
- [8] 牧田 孝嗣, 神原 誠之, 横矢 直和: "ネットワーク共有データベースとハイブリッドP2Pを利用したウェアラブルシステムユーザへの注釈付け," 日本バーチャルリアリティ学会論文誌, Vol. 13, No. 2, pp. 171-181, 2008.