

デジタル地球における NUI を用いた主要撮影地域探検システム

橋本 和哉

熊野 雅仁

木村 昌弘

龍谷大学理工学部電子情報学科

1 はじめに

写真は魅力的な対象に出くわしたとき撮影される傾向がある。近年、ソーシャルメディアに蓄積された撮影位置・撮影時間・撮影者情報が付随した大量の写真を集約することで、魅力的な地域や旬シーズンを抽出し、可視化する研究 [1] や、Google Earth などのデジタル地球に没入し、魅力的な場所を探索する環境 [2] などが注目されている。観光への応用に向け、これらの研究をより発展させるため、本研究では、デジタル地球上に没入し、自然かつ効率良く観光地を疑似体験しながら各地の魅力の発掘や旅行プラン生成の支援環境実現を目指している。そこで、ユーザ自身の行う手や足のジェスチャ認識や音声認識に基づく NUI (Natural User Interface) を構築し、デジタル地球上に可視化された主要撮影スポット上の写真情報を効率良く閲覧しながら地域の魅力を調べ、旅行プラン生成支援への応用も期待される主要撮影地域探検システムを開発している。本研究では、実装した試作システムの動作検証を行う。

2 提案システム

2.1 提案する NUI

デジタル地球内の主要撮影地域を巡り、写真を閲覧しユーザ好みの旅行プランを練り上げるには、仮想空間内の移動（前後進・加減速・停止・方向転換）や数多く存在する写真の選択・切替など多種の操作をシームレスに効率良く行いつつ負担の少ない環境が望まれる。

本研究では、ユーザがデジタル地球上を探検する時間が長期化する場合を考慮し、ユーザへの負担を減らすため、立った状態での使用を控え、負担が少ない座った姿勢を基本姿勢 (図 1(a)) とする。また、常識を逸脱せず直観的に理解でき、一度の説明で容易に習得し得るジェスチャによる操作法を考える。ただし、写真操作は、移動中に行うとユーザの意識散漫や混乱を招き、負担になると考え、空間内での静止時のみ可能とする。つまり提案法は移動モードと静止モードを設ける。特に、写真数が多い場合に写真の切替操作が頻発する負担を軽減する観点から、写真選択や切替操作に音声認識を用いる。つまり、本研究では、ジェスチャと音声の認識ならびに移動モードと静止モードをシームレスにつないで効率的な操作を可能とする NUI を提案する。

頭部姿勢に追従した没入空間の視野表示機能

本研究では、頭部姿勢センサを装着した没入型 HMD

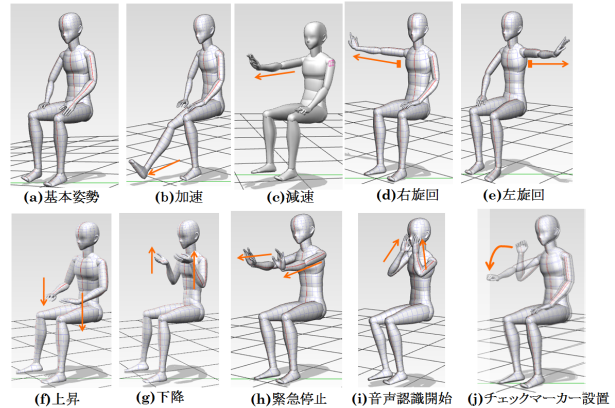


図 1: 認識されるジェスチャ

を用いてデジタル地球内を見るシステムを実装した。仮想空間内の視野がユーザの頭部の動きに追従して動くため、現在地点からデジタル地球内を見まわすことができる。この環境に基づき各種操作機能を説明する。ジェスチャに基づく空間内の移動操作機能

まず、初期状態で基本姿勢 (図 1(a)) であれば、静止モードとなっている。本研究では、ジェスチャ操作によるユーザの負担を軽減するため、手や足を上げ続けるような操作法を極力避けて、一瞬のジェスチャで操作を行う手法を考える。以下は、すべて一瞬だけジェスチャを行う手法である。例えば、右もしくは左足を前方に一瞬出す (図 1(b)) と移動モードとなり、基本姿勢に戻っても移動し続ける。もう一度出すと、現在の移動方向でスピードを増す。一方、片手を前方に上げる (図 1(c)) たびに一段階減速する。また、モードに関係なく右方向に右手を上げると右旋回 (図 1(d)) し、左方向に左手を上げると左旋回 (図 1(e)) する。進行方向変更の際、ジェスチャを行う時間を短縮するために体から手を遠くに上げるほど大きく旋回を行うようにしている。ところで、実世界で空を飛ぶことは容易ではないが、提案システムでは、静止モードで、手を押し下げると高度が上昇 (図 1(f)) し、手を押し上げると高度が下降 (図 1(g)) し、前進と同じく基本姿勢に戻っても高度を維持し続ける。非現実的ではあるが、これにより、地上視と鳥瞰視をシームレスに結ぶため、視覚的な迷子を防ぐこと、及び周辺の観光スポット発見に貢献する。また、両手を前方に上げる (図 1(h)) ことにより前進及び高度変更を緊急 (完全) 停止する。このとき、静止モードとなる。つまり、一瞬のジェスチャでデジタル地球内の全ての動きを停止することができる。

音声とジェスチャによる写真群の選択・切替操作機能

提案システムは、大量の Geo-tag 写真を集約することで抽出された主要撮影地域の写真群を閲覧することにより各地の魅力の発見を促し、旅行プラン生成を支援する。ただし、各主要撮影地域には数多くの写真群が含まれるだけでなく、季節や時間帯ごとに主要な撮影対象が変わり得る。効率良く写真を閲覧するため、試作システムの音声認識では「春、夏、秋、冬、朝、昼、夜」を認識し、各主要撮影地域の写真群を四季および三つの時間帯を選択して切り替える機能を実装した。ただし、常に音声認識を行うと、認識する音声以外の自由な発話を奪い負担にもなるため、図 1(i) のように、口前に両手をだすと音声認識を開始し「終了」の言葉を認識すると音声認識を終了するよう実装した。また、図 1(j) のようなジェスチャにより、ユーザがデジタル地球内で気に入った場所に座標、時間情報が付随するチェックマーカをいくつでも設置できるようにした。これにより、デジタル地球内の探検後、マーカを置いた位置・時間情報を Google Earth 上で確認することで、ユーザは各地の巡り方を吟味することができる。

2.2 システム概要

本研究では、GoogleEarth 上で画像を提示可能にするため、写真共有サイト Flickr から Geo-tag 付き写真を収集し、KML 形式で記述されたデータセットを作成する。また、ジェスチャ認識には Microsoft の Kinect を用い、音声認識には、Microsoft の Speech Platform を用いて実装した。

3 実験

3.1 実験設定

本研究では、試作システムの動作検証を行うため、日本有数の観光地、京都の主要撮影スポットを巡る実験を行った。写真共有サイト Flickr から、2012 年 1 月 1 日～12 月 31 日までの WoEID を Kyoto と指定し、収集した 52398 枚の Geo-tag 付写真を対象として、主要撮影スポット 27 地域を抽出 [1] し、主要撮影スポットごとに写真群を春:3 月 1 日～5 月 31 日、夏:6 月 1 日～8 月 31 日、秋:9 月 1 日～11 月 30 日、冬:12 月 1 日～2 月 28 日の四季に分割し、さらに各季節を朝:5 時台～11 時台、昼:12 時台～18 時台、夜:19 時台～4 時台の時間帯に分類した。

3.2 実験結果

京都の様々な主要撮影スポットの探索を行い、動作検証を行った。図 2 に実際にデジタル地球を探検している様子を示す。図 2 の (a) や (b) の左側に Kinect により認識されたユーザの骨格モデル、図 2 の (a) や (b) の右側に HMD に表示された視野を示す。視野中、近隣や遠方 (図 2(b) 右側の赤丸) に浮かぶ画像群は各主要撮影スポットに対応する写真群である。上空 (図 2(b)) では、写真群の位置から魅力的な地域の位置や現在位置からの距離感が把握でき、地上 (図 2(b)) では、写真群



図 2: デジタル地球内の移動操作 (左:地上、右:上空)

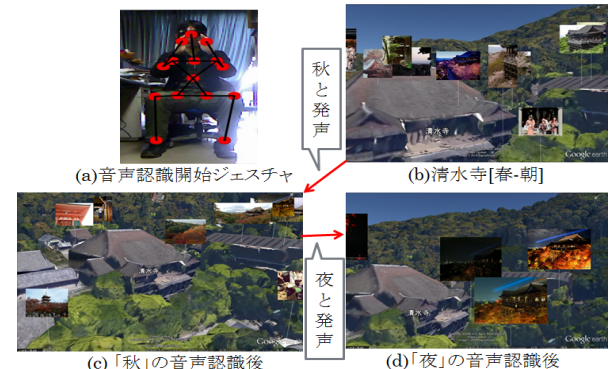


図 3: 音声とジェスチャの認識による写真群の時間情報変更例

に近づくとも写真の内容から地域の魅力を詳しく把握し得ることがわかる。また、図 3 は主要撮影地域の清水寺に到達した例である。図 3(a) に図 1(i) のジェスチャ認識と音声認識を行い写真群の季節や時間帯 (時間情報) を変更する例を示す。初期状態 (図 3(b)) では春の日の出 (朝) における桜が写った清水寺の写真が見えていたが、図 3(a) のジェスチャで音声認識を開始し、「秋」の発声の後、図 3(c) となり、「夜」と発声した後、図 3(d) のように紅葉の夜景が写った写真群に変化した。このように、ジェスチャと音声の認識が良好に動作しており、撮影スポットを巡り、現地を疑似体験しながらシームレスに写真群切り替え操作を行い、主要撮影スポットにおける魅力の発掘を支援していることがわかる。

4 まとめ

本研究では、ユーザ自身の行う手や足のジェスチャ認識や音声認識などに基づくマルチモーダルな NUI (Natural User Interface) を構築し、デジタル地球上に可視化された主要撮影スポット上の写真情報を効率良く閲覧しながら地域の魅力を調べ得る、主要撮影地域探検システムを提案した。実験により、提案システムによってユーザ自身の四肢、頭部姿勢、声がデジタル地球上の各操作に直結し自然かつ効率よくデジタル地球の探検が行えた。今後は、疑似体験や旅行プランの生成をより良く支援するための機能拡張を行う予定である。

参考文献

- [1] 熊野, 岩瀬, 小関, 小野, 木村: 集合知に基づいたポピュラー撮影スポットに関する旬シーズンの可視化, 芸術科学会論文誌, Vol.13, No.4, pp.218-228, 2014.
- [2] Kisilevich, S., Mansmann, F., Bak, P., Tchaikin, A., Keim, D., Where would you go on your next vacation? A framework for visual exploration of attractive places, GEOPROCESSING '10