2視点における音響カメラ画像を用いた 水中物体の特徴点の3次元計測

○郭承澈(東京大学) 池勇勳(東京大学) 山下淳(東京大学) 淺間一(東京大学)

1. 序論

福島原子力発電所の原子炉のような災害現場や人の 入れない水中環境では、ロボットによる状況調査が多 く行われている.それらの状況調査を行うためには、水 中における物体認識が極めて重要である.特に、調査 対象の3次元情報を取得することは、水中環境を把握 する上で最も役に立つ方法である.

水中に存在する調査対象の3次元形状を復元するに は、水中において何らかの情報を取得する必要がある. その方法の1つに、光学カメラから得られる画像が用い られる. Shibataら[1]は光学カメラを用いたスケール 復元が可能な水中 Structure from Motion を提案し、水 中物体の3次元測定を行った.光学カメラからは解像 度の高い画像が得られるため、水中環境の把握が容易 である.しかし、光学カメラは透明な水中環境に適し ているセンサであり、濁った水中では使用不可能とい う限界がある.また、水中物体認識を行うために、レー ザセンサ[2]も用いられている.しかし、レーザや電磁 波のような周波数の高い信号は、水中において減衰が 早く進行される問題が挙げられる.

したがって,水中環境情報を取得するためには, 一般的に超音波センサが有効である.特に,近年は DIDSON (*Dual-Frequency IDentification SONar*) [3] や ARIS (Adaptive Resolution Imaging Sonar) [4] などの ような音響カメラが活躍している.音響カメラは,従 来の超音波センサでは不可能であった3次元空間のセ ンシングが可能であり,なおかつ解像度の高い画像を 出力することができる.しかし,図1に示すように音 響カメラ画像は光学カメラ画像と相違があり,音響カ メラ画像を解析するモデルは存在しない.Yuら[5]は, 予め撮影した音響カメラ画像を用いて学習データを構 築し,それらのデータに基づいて調査対象を区別する システムを提案した.しかし,このシステムは学習デー タに大きく依存され,未知の調査対象に対しては認識 が不可能である問題点がある.

以上の理由から、本研究では音響カメラ画像そのも のを解析し、2視点における音響カメラ画像を用いて調 査対象の特徴点の3次元計測を行うモデルを提案する.

2. 音響カメラの基本原理

2.1 音響カメラのプロジェクションモデル

音響カメラ画像を生成するために、音響カメラはある 3次元空間の範囲で超音波を発する.その範囲は、図2 に示すように最大測定距離 R_{cam} ,方位角 θ_{cam} ,そし て仰角 ϕ_{cam} によって定まる.これらのパラメータは音 響カメラの仕様によるものであり、(R_{cam} , θ_{cam} , ϕ_{cam}) で定まった範囲をセンシング領域とする.音響カメラ から発された超音波はセンシング領域を進行し、物体 にあたってから反射される.超音波は進行及び反射に よってその強度が弱まる.反射された超音波は音響カ メラに戻って来て,測定距離 R,方位角 θ ,そして弱 まった反射強度 I の関数として処理される.

2.2 音響カメラ画像生成における幾何学的モデル

図2に示すようにセンシング領域にある調査対象の 点 Pは、音響カメラを基準としたモデル座標(R_P , θ_P , ϕ_P)を有する.しかし、音響カメラ画像に影響を与え る座標は(R_P , θ_P)であり、 ϕ_P の値は無関係である. この現象は、音響カメラ画像が測定距離 R,方位角 θ , そして反射強度 Iの関数として処理されるからである. 音響カメラ画像における画素の座標値は測定距離 R, 方位角 θ で定まり、それらで定まった画素には反射強 度 Iに基づいた値が入力される.したがって、仰角の 値は異なっても測定距離及び方位角の値が同じである 調査対象は、音響カメラ画像において同じ画素に出力 される.この場合、各反射強度の合計が出力される.

3. 3次元復元モデル

2 章で述べたように,音響カメラ画像は測定距離 *R*, 方位角 θ,そして反射強度 *I*の関数として処理される. つまり,音響カメラ画像より調査対象の 3 次元形状を



(a) 音響カメラ画像

(b) 光学カメラ画像

図1 角柱を撮影した音響カメラ画像及び光学カメラ画像



図2 音響カメラのプロジェクションモデル

復元することは不可能である.これは、仰角φの値は 算出できないからである.しかし、位置関係が分かる 2 視点を用いることによって調査対象の3次元情報は 一意に決まる.本研究では、位置関係が既知である2 つの視点を用いて、調査対象の特徴点の3次元情報を 取得するモデルを提案する.

3.1 特徴点の抽出

まず,各視点における音響カメラ画像より特徴点を 抽出する.特徴点とは,音響カメラ画像において認識 可能な調査対象を指しており,反射強度が急変する頂 点や材料の異なるものなどが考えられる.

3.2 特徴点の存在可能な候補点の抽出

図2に示す特徴点 Pの正確な3次元座標を求めるこ とは不可能であるが,音響カメラ画像における画素座 標より測定距離 R_P ,方位角 θ_P を求めることはでき る.仰角 ϕ_P の値に関しては音響カメラの仕様による範 囲で表現すると,特徴点 Pの存在可能な場所を求める ことができる.視点1において特徴点 Pの存在可能な 座標を ${}^{1}P_i = [{}^{1}R^P, {}^{1}\theta^P, {}^{1}\phi_i^P]^T$ とする.ただし,iは特 徴点 Pの存在可能な座標のインデックスであり, ${}^{1}\phi_i^P$ は,

$$0 \le {}^1 \phi_i^P \le \phi_{\text{cam}} \tag{1}$$

である. 特徴点 *P* に対する候補点の集合を ¹**P** とする と, ¹**P** は

$${}^{1}\mathbf{P} = [{}^{1}P_{1}\cdots{}^{1}P_{i}\cdots{}^{1}P_{I}]^{T}, \qquad (2)$$

と表され、各候補点は

$${}^{1}P_{1} = [{}^{1}R^{P}, {}^{1}\theta^{P}, {}^{1}\phi_{1}^{P}]^{T},$$

$$\vdots$$

$${}^{1}P_{i} = [{}^{1}R^{P}, {}^{1}\theta^{P}, {}^{1}\phi_{i}^{P}]^{T},$$

$$\vdots$$

$${}^{1}P_{I} = [{}^{1}R^{P}, {}^{1}\theta^{P}, {}^{1}\phi_{I}^{P}]^{T},$$

(3)

として表される.ここで,*I*は視点1における特徴点 *P*の候補点の数である.同様に,視点2において特徴 点*P*の存在可能な座標のインデックスを*j*とすると, その座標は ${}^{2}P_{j} = [{}^{2}R^{P}, {}^{2}\theta^{P}, {}^{2}\phi_{j}^{P}]^{T}$ と書ける.それら 候補点の集合を ${}^{2}\mathbf{P}$ とすると, ${}^{2}\mathbf{P}$ は

$${}^{2}\mathbf{P} = [{}^{2}P_{1}\cdots{}^{2}P_{j}\cdots{}^{2}P_{J}]^{T}, \tag{4}$$

と表され、各候補点は

$${}^{2}P_{1} = [{}^{2}R^{P}, {}^{2}\theta^{P}, {}^{2}\phi_{1}^{P}]^{T},$$

$$\vdots$$

$${}^{2}P_{j} = [{}^{2}R^{P}, {}^{2}\theta^{P}, {}^{2}\phi_{j}^{P}]^{T},$$

$$\vdots$$

$${}^{2}P_{J} = [{}^{2}R^{P}, {}^{2}\theta^{P}, {}^{2}\phi_{J}^{P}]^{T},$$

(5)

として表される. ただし, *J* は視点2における特徴点 *P* の候補点の数である.

3.3 3次元座標の算出

前述のように各視点における音響カメラ画像からは 特徴点の3次元座標の候補点しか求められないが,位 置関係が分かっている2視点を用いると,図3に示す ように特徴点の3次元座標は一意に決まる.各視点に おいて候補点が存在する2つの円弧は1点で交わるか らである.

本研究で提案するモデルでは、各視点における候補 点間の距離が最小となるときの候補点を採用し、それ らの各座標値の平均値を特徴点の3次元座標とする.各 視点における候補点間の距離が最小になるときを採用 することで、実装上2つの円弧が1点で交わらない場 合に対しても対応可能であり、なおかつ理論上交わる であろう点の座標と近接な座標値が得られるからであ る.その祭に、極座標系で表現されている各候補点を デカルト座標系に変換する.デカルト座標系に変換し た視点1における特徴点 P の候補点を,

$${}^{1}P_{i} = [{}^{1}x_{i}^{P}, {}^{1}y_{i}^{P}, {}^{1}z_{i}^{P}]^{T},$$
(6)

とする.そして,視点2における特徴点Pの候補点を,

$${}^{2}P_{j} = [{}^{2}x_{j}^{P}, {}^{2}y_{j}^{P}, {}^{2}z_{j}^{P}]^{T},$$
(7)

とすると候補点間の距離 l(i, j) は,

$$\begin{split} l(i,j) &= \\ \sqrt{({}^{1}x_{i}^{P}-{}^{2}x_{j}^{P})^{2}+({}^{1}y_{i}^{P}-{}^{2}y_{j}^{P})^{2}+({}^{1}z_{i}^{P}-{}^{2}z_{j}^{P})^{2}}, \end{split}$$

となる.そして候補点間の距離を最小とする各視点に おけるインデックスを (i_{\min}, j_{\min})とすると,

$$(i_{\min}, j_{\min}) = \underset{1 \le i \le I, 1 \le j \le J}{\operatorname{arg min}} l(i, j), \tag{9}$$

となる. また, インデックスを (*i*_{min}, *j*_{min}) による候補 点の各座標値の平均をとると, 特徴点 *P* の 3 次元座 標は,

$$P = \begin{pmatrix} \frac{(^{1}x_{i_{\min}}^{P} + ^{2}x_{j_{\min}}^{P})}{2} \\ \frac{(^{1}y_{i_{\min}}^{P} + ^{2}y_{j_{\min}}^{P})}{2} \\ \frac{(^{1}z_{i_{\min}}^{P} + ^{2}z_{j_{\min}}^{P})}{2} \end{pmatrix},$$
(10)

となる.

以上より、2視点における音響カメラ画像を用いることで特徴点 Pの3次元座標を求めることができる. 複数の特徴点を利用することで、その調査対象の3次元形状を復元することができる.

4. 3次元復元実験

本章では第3章で述べたモデルを用いて,調査対象 の3次元形状を復元する実験について述べる.各視点 の音響カメラ画像における複数の特徴点を利用するこ とによって,調査対象の形状を復元することができる. ここでは,我々が開発したシミュレータ[6]によるシ



図3 2視点を用いた3次元座標の復元

ミュレーション実験を行う. また, 調査対象は図 1(b) に示すように複数の特徴点が抽出できる角柱を用いる ことにした.

図4に角柱を撮影した音響カメラ画像を示す.図4(a) は視点1における音響カメラ画像であり,図4(b)は視 点2における音響カメラ画像である.また,図中に示 している点は特徴点を表していると同時に2つの視点 に対する特徴点間の対応関係を示している.これらの 特徴点の3次元座標を復元することで調査対象の形状 の把握ができる.

本研究では, Sound Metrics 社の ARIS EXPLORER 3000 という音響カメラの仕様に基づいてシミュレー ション実験を行った.その仕様は表1に示されている.

まず,視点1の音響カメラ画像における各特徴点の 画素座標を抽出する.画素座標により測定距離及び方 位角 (${}^{1}R^{N}, {}^{1}\theta^{N}$)が算出される. N は,それぞれの特 徴点に対する上添字とする.仰角に関しては,表1に 示している仕様を用いると,

> ${}^{1}P^{i} = [{}^{1}R^{P}, {}^{1}\theta^{P}, {}^{1}\phi^{P}_{i}]^{T},$ \vdots ${}^{1}U^{i} = [{}^{1}R^{U}, {}^{1}\theta^{U}, {}^{1}\phi^{U}_{i}]^{T},$ (11)

と表すことができる.ただし, $i = 0, 1, \dots, 1400$ であり、 ${}^{1}\phi_{i}^{N} = \phi_{cam} - 0.01 \times i$ [deg] である.次に、視点2の音響カメラ画像における各特徴点の画素座標を抽出した後、上述と同様に画素座標により測定距離及び方位角 (${}^{2}R^{N}, {}^{2}\theta^{N}$)を算出する.仰角に関しては、視点1の場合と同様に適用すると、

$${}^{2}P^{j} = [{}^{2}R^{P}, {}^{2}\theta^{P}, {}^{2}\phi_{j}^{P}]^{T},$$

$$\vdots$$

$${}^{2}U^{j} = [{}^{2}R^{U}, {}^{2}\theta^{U}, {}^{2}\phi_{j}^{U}]^{T},$$

(12)

と表すことができる.ただし, $j = 0, 1, \cdots, 1400$ であ り, ${}^{2}\phi_{j}^{N} = \phi_{cam} - 0.01 \times j$ [deg] である.極座標で表現 されている各特徴点の座標をデカルト座標に変換した 後,式(8),(9),(10)を用いて各特徴点の3次元座標を 求める.上述より求めた特徴点の3次元座標の計測値, そしてそれらをプロットした結果を表2,3及び図5に 示す.図5に示すように,本実験では,調査対象であ る角柱の頂点6箇所の3次元情報を復元した.赤丸い 点は各頂点の真値を表しており,青いダイアモンド点 は各頂点の計測した座標値を表している.



表1 ARIS EXPLORER 3000 の主な仕様 [7]

最大測定距離 B	5 m
取入例足距離 n _{cam}	5 111
方位角 $ heta_{cam}$	32 deg
仰角 $\phi_{ m cam}$	14 deg
ビームの数	128
ビーム幅	0.25 deg



図5 各特徴点の真値及び計測値

5. 考察

シミュレーションの結果から,理論的には特徴点の3 次元座標を求め,調査対象の形状復元に適用可能である ことが確認できた.しかし,表3をみると,0.02~0.03 m 程度の誤差が生じている.この原因として考えられる のが,音響カメラ画像座標の精度である.連続的であ る実空間の座標は,離散的である音響カメラ画像画素 に出力される.ピクセル単位の整数値で取得した画像 座標を処理したため,真の3次元座標に対して誤差が 生じたと考えられる.

また, 誤差が生じた理由としてビーム幅が考えられ る. 図 6 に示すように, ARIS EXPLORER 3000 は方 位角 32 deg を, 幅が 0.25 deg であるビーム 128 個に 分けて, 各ビーム毎に処理を行う. そのため, ビーム 幅 0.25 deg から生じてしまう誤差は音響カメラの特性 上, 避けることができない. 本研究のシミュレータで は, 音響カメラから水中物体までの距離を 6 m 前後に 設定しており, ビーム幅によって生じる誤差は 0.026 m 前後である. そのため, 表 3 に示しているような誤差 が発生したと考えられる.

6. 結論

本研究では2視点における音響カメラ画像を用いて 水中物体の特徴点の3次元情報を取得するモデルを提 案した.このモデルにより従来ではできなかった音響 カメラ画像そのものの解析ができた.水中物体の特徴 点の3次元計測ができることを確認し,水中環境の把 握に貢献できることを明らかにした.

実画像を用いた実測実験は今後の課題である.



(a) ARIS EXPLORER 5000 の ビーム構成

誤差

図 6 ARIS EXPLORER 3000 のビーム

表 2 特徴点の真値座標及び計測座標

特徵占	直值	計測座標 [m]
131-2/11	八四元 (m)	II MALIN [III]
P	[5.500, 5.300, 1.000]	[5.500, 5.294, 1.043]
Q	[5.500, 4.700, 1.000]	[5.502, 4.723, 1.031]
R	[4.500, 4.700, 1.000]	[4.520, 4.700, 1.049]
S	[4.500, 5.300, 1.000]	[4.516, 5.301, 1.057]
T	[4.500, 5.300, 0.000]	[4.516, 5.323, 0.050]
U	[4.500, 4.700, 0.000]	[4.511, 4.707, 0.037]

表3 真値に対する計測座標の誤差

特徴点	RMSE [m]	特徴点	RMSE [m]
P	0.025	S	0.022
Q	0.030	T	0.034
R	0.033	U	0.023

謝辞

本研究の一部は,総合科学技術・イノベーション会 議により制度設計された革新的研究開発促進プログラ ム (ImPACT)「タフ・ロボティクス・チャレンジ」の 援助を受けた.

また、本研究で実施した実験は極東建設、東陽テク ニカ及び日創建の支援によって行われた.

参考文献

- A. Shibata, H. Fujii, A. Yamashita, and H. Asama, "Scale-reconstructable Structure from Motion Using Refraction with a Single Camera," *Proceedings of the* 2015 IEEE International Conference on Robotics and Automation, pp. 5239-5244, 2015.
- [2] Y. Chang, F. Peng, L. Luo, Y. Zhang, "Laser Imaging for the Underwater Object and the Image Segmentation Based on Fractal," *Proceeding of the 2003 International Symposium on Multispectral Image Processing and Pattern Recognition*, pp. 668-671, 2003.
- [3] E. Belcher, W. Hanot, and J. Burch, "Dual-frequency Identification Sonar (DIDSON)," *Proceedings of the 2002 IEEE International Symposium on Underwater Technology*, pp. 187-192, 2002.
- [4] "Aris," 2015, retrieved June 05, 2015, from http://www.soundmetrics.com/Products/ARIS-Sonars/ARIS-Explorer-3000.
- [5] S. C. Yu, J. H. Kim, T. Zhu, and D. J. Kang, "Development of 3D Image Sonar Based Object Recognition for Underwater Vehicle," *Proceedings of the 2012 International Offshore and Polar Engineering Conference*, pp. 502-507, 2012.
- [6] S. Kwak, Y. Ji, A. Yamashita, and H. Asama, "Development of Acoustic Camera-Imaging Simulator Based on Novel Model," *Proceedings of the 2015 IEEE International Conference on Environment and Electrical Engineering*, pp. 1719-1724, 2015.
- [7] "Specifications of ARIS EXPLORER 3000,"
 2015, retrieved June 05, 2015, from http://www.soundmetrics.com/Products/ARIS-Sonars/ARIS-Explorer-3000/ARIS-3000-Product-Specs-English.