

## 画像解析による水中画像からの魚類計測

## Fish Categorization and Counting Method from Underwater Video and Images using Image Analysis

堤 富士雄<sup>\*1</sup>  
Fujio Tsutsumi立田 穰<sup>\*1</sup>  
Yutaka Tateda<sup>\*1</sup> 一般財団法人 電力中央研究所  
Central Research Institute of Electric Power Industry

In order to evaluate global warming effect to coastal ecosystems, investigations of the fish key species composition change will be useful. In this report, to support the fish investigation methodology, an automatic categorization and counting technique using image processing is newly proposed. The proposal technique uses video data to detect moving fish, and uses high resolution images to determine fish category. The video data is captured by a video camera, and the high resolution images are captured by a digital camera using a time lapse shooting feature. The video camera and the digital camera are juxtaposed underwater. Our technique detects moving fishes from the video data by a background subtraction. Next, the correspondence fish position at an image is calculated from the detected position at a video frame. Finally fish is categorized from the detailed fish image by a machine learning technique. We evaluated the proposal technique by the underwater 26 minutes video recorded in Ishigaki Island and the 157 still pictures captured every 10 seconds simultaneously. As a result of the experiment, the fish detection from the video is 98% of precision, 88% of recall, and 71% of matching accuracy. Fish recognition is 94% of accuracy by the elementary 2 classes recognition (fish or not). In order to perform all the processes automatically, 70% of matching accuracy is a bottleneck. However, it is useful from the ability to automatically find out many fishes to support human's investigation.

## 1. はじめに

気候変動や将来の地球温暖化が、沿岸生態系の多様性や水産資源に及ぼす影響を明らかにするには、魚類の生息量調査が必要であり、網による捕獲やダイバーによる目視調査が行われている[西田 2005][亀井 2009][桑原 2010]。これに対し、従来の調査法を補うため、環境影響が少なく、長時間の計測が可能な、画像機器による水中撮影への期待が高まっている。しかし、撮影した動画や静止画をもとに魚種や生息量を把握するには、人手による長時間の解析を要するという欠点があり、記録画像から魚種や生息量を自動計測する手法が求められている。そこで本稿では、水中で撮影した動画と静止画とを組み合わせ、自動的に魚の種類や量を計測する手法を提案する。組み合わせる理由は動画と静止画の利点を生かし、それぞれの欠点を補うためである。

## 2. ビデオカメラとスチルカメラの連携上の課題

## 2.1 ビデオカメラとスチルカメラを連携した魚類撮影

提案手法は、動画と静止画の両方を用いて、魚影の検出・判別を行う点に特徴がある。動画は時間単位の画像枚数が多く、移動する対象を画像処理で抽出するのに適している。しかし解像度に制限があり、撮影された魚影はカメラから離れると小さくかつ不鮮明となるため、魚種判別は困難である。一方で、静止画像撮影用のスチルカメラは高解像度であり、魚種判別に適している。ただしスチルカメラは時間単位の撮影枚数に制限があり、高速長時間撮影と高解像度撮影を両立する映像記録装置は普及していない。そこで提案手法では、魚影の位置検出は動画で、魚影の種別判定は静止画像で行うように、役割分担をし

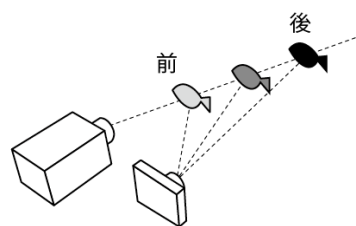


図 1: 視差ズレ

つつ連携することで目的を達成する。

## 2.2 連携上の時間的・空間的ズレ

2 台の撮像装置を用いる提案手法では、動画と静止画を対応付ける必要がある。異なる撮像系で撮影した動画と静止画には、時間的なズレと空間的なズレがあり、その補正が必要となる。

## (1) 時間的ズレ

魚は移動するため、動画と静止画の時刻が合っていないと連携は困難になる。電子的に連携していないカメラの時計に 1 秒以内の同期精度を確保するのは困難である。そのため撮影後に後述するノウハウにより、動画と静止画の時刻を数フレーム以内(1フレームは 33 ミリ秒)に合わせるが、それでも誤差を無くすことはできない。この残された誤差は、主にスチルカメラのメカニズムとタイムスタンプ(撮影時刻情報)の精度に原因がある。

## (2) 空間的ズレ

## (2)-a 視差ズレ

視差ズレは、人間の右目と左目のように、2台のカメラ間(レンズの光軸間)に距離があるために、像がずれるという現象である(図 1)。このズレは2台のカメラの間が離れば離れるほど大きくなる。またズレはカメラの近くほど大きく、遠方では小さい。提案手法では、動画から得た魚影位置から、静止画面上の位置を

連絡先: 堤 富士雄, 電力中央研究所 システム技術研究所,  
〒201-8511 東京都狹江市岩戸北 2-11-1, 03-3480-2111,  
tutumi@criepi.denken.or.jp

推定する。しかし、奥行き情報が分からない状態では、解を求めるために必要な情報が不足している不良設定問題である。次章で述べる提案手法では部分画像照合により、この問題に対策する。

### (2)-b 撮像機構・過程の違い

対応付け問題を困難にする主要因ではないが、視差ズレ以外にも、レンズの画角、歪、焦点距離などのカメラパラメータ、解像度、露出、ホワイトバランス、インターレースの有無も手法の効果に影響する。

## 3. ビデオカメラとスチルカメラを連携した魚種定量手法

提案手法は以下の3つのブロックで構成されている。

- (1) 魚影抽出ブロック(動画)
- (2) 魚影対応付けブロック(動画→静止画)
- (3) 魚影判定ブロック(静止画)

魚影抽出ブロックでは、撮影された動画から動き情報を用いて魚影を抽出する。次に魚影対応付けブロックでは、動画から抽出された魚影の位置を、静止画面上に対応付ける。最後に魚影判定ブロックで、静止画像上の魚種を自動判定する。

### 3.1 魚影抽出ブロック

提案手法では、動体検出手法の一つであり、当所で既開発の背景差分を用いた魚影抽出手法[堤 2005]を応用する。この手法では、水面や水中など光のゆらぎが大きい環境で効果的に機能することを目的として、時間方向の色値の分散を考慮した背景データを作成する。具体的には、画面内の座標それぞれに対して、画素値の時間方向の平均と標準偏差を、一定時間の範囲で計算する。つまり画面サイズが横  $w$  画素、縦  $h$  画素の場合、 $w \times h$  個の画素それぞれに対して時間平均  $\mu$  と標準偏差  $\sigma$  を計算する。RGB 表色系の場合は、R, G, B それぞれに対し平均と標準偏差を求め、これらを背景データとする。この平均と標準偏差は、ある時間範囲内での画素値の変動を表す。

差分の判定は、グレースケールで輝度のみを扱うとすると、ある時刻での、ある座標における輝度  $g$  に対し、 $[\mu - \beta \sigma, \mu + \alpha \sigma]$  の区間(シグマバンド)外にある場合に差分が生じたと判定する。ここで  $\alpha$  と  $\beta$  は判定を調整するパラメータである(4章の水中動画の場合は  $\alpha = \beta = 4.0$  とした)。RGB 表色系でのカラーの場合は、R, G, B それぞれのシグマバンドのいずれかの外にある場合に差分が生じたと判定する。

差分画素の抽出後は、ノイズ除去、近傍画素の結合および、フレーム間の領域追跡、移動速度に関するフィルタリングを行って魚影を抽出する。

### 3.2 魚影対応付けブロック

魚影対応付けブロックは、2つのステップから構成される。一つは、前節の手法で抽出された魚影座標を、静止画上の座標に変換する座標変換ステップである。もう一つのステップは、その変換された座標を基準に静止画像内で確からしい魚影位置を探索する魚影探索ステップである。

#### (1) 座標変換ステップ

座標変換ステップでは、複数のカメラ座標の変換方法として最も単純な線形変換を用いる。図 2 を用いて変換方法を説明する。ユーザがあらかじめ動画と静止画との間で、対応付け可能な3点を基準点として指定済みとする。基準点には、海底のサンゴや、ダイバーの手足、特徴的な色や模様の魚などを用い

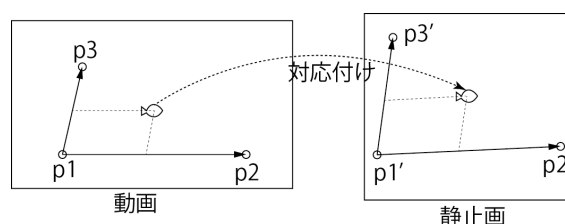


図 2: 動画から静止画への座標変換

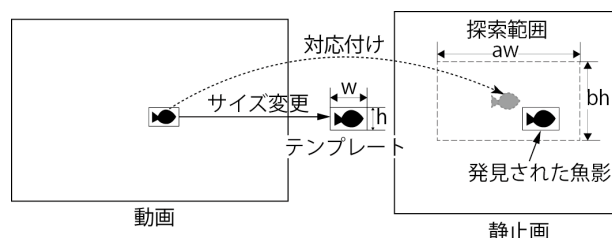


図 3: テンプレート照合による魚影探索

る。また格子状の模様が描かれた板をカメラの前に掲示して、目印にすることも可能である。

動画上での基準点をそれぞれ  $p1, p2, p3$ 、静止画上の対応する基準点を  $p1', p2', p3'$  とする。これら3点のうち任意の1点、例えば  $p1$  を原点として、 $p1$  から  $p2$  へのベクトルと、 $p1$  から  $p3$  へのベクトルをそれぞれ単位ベクトルとする斜交座標系  $A_v$  を設定する。同様に動画上で原点とした点に対応する  $p1'$  を原点として、 $p1'$  から  $p2'$  へのベクトルと、 $p1'$  から  $p3'$  へのベクトルを単位ベクトルとする斜交座標系  $A_s$  を設定する。魚影位置の変換は、斜交座標系  $A_v$  上での座標を、そのまま斜交座標系  $A_s$  での座標とすることで行う。

基準点とした3点は水中の空間内の一つの平面を規定し、変換はこの平面内での対応付けを意味する。原理的には、この平面内を魚が移動する範囲では、対応付けは正確に行われる。しかし前章で述べた視差ズレの影響で、魚影がこの平面から離れると、ビデオカメラで同じ場所に見える魚影が、スチルカメラでは異なる位置に見える。

#### (2) 魚影探索ステップ

魚影探索ステップでは、静止画上で魚影を探索する。これは視差ズレにより、正しい座標を示していないとしても、同じ時刻の画像であれば、近傍に魚影が見当たる筈との仮定に基づく。

図 3 を用いて探索方法を説明する。まず魚影抽出ブロックで得られた、動画上の長方形の魚影領域の画像を、静止画像上のサイズに合わせて線形変換により伸縮し、これをテンプレートとする。テンプレートの横幅を  $w$ 、縦の長さを  $h$  とした場合、座標変換ステップで得られた静止画上の対応点を中心に、 $aw$  の横幅、 $bh$  の縦の長さの探索範囲を設定する。ここで  $a, b$  は探索範囲の広さを調整するパラメータである(後述の評価実験では  $a=5, b=3$  とした)。探索範囲内をスキャンし、テンプレートと最も類似度の高い位置を、魚影位置とする。テンプレートサイズにより探索範囲を変える理由は、カメラに近いほど視差が大きくなることを考慮し、大きく見える場合は近い位置にあると見なし、探索範囲を広くするためである。

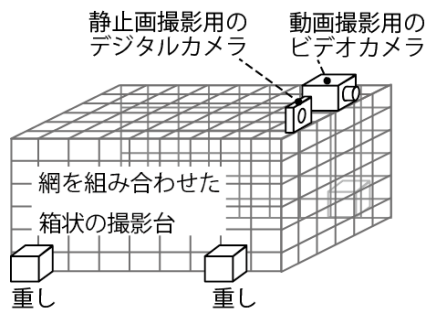


図 4: 水中撮影機材

類似度の判定には、グレースケールでの正規化相互相関を用いた。グレースケールとした理由は、動画と静止画でホワイトバランスの違いにより色が異なるためである。正規化相互相関を用いた理由は最も代表的かつ実用的な、テンプレート照合方式であるためである。水中の魚影映像に合わせた照合方式の改善は今後の課題である。

### 3.3 魚影判定ブロック

自動判定は、画像から特徴量を計算し、機械学習することで、自動判定を行うための判定基準を作成する。判定基準の作成には、魚影画像と、人間がその魚影を判定した結果(ラベルと呼ぶ)の対を集める必要がある。以降これを画像データベースと呼ぶ。計算機が判定基準を作成する方法は、我々が開発した既存の手法の応用である[中島 2005]。プロトタイプでの機械学習には、パターン識別器として一般的な SVM (Support Vector Machine) を用いた。一方特徴量に関しては、現状は、画像データベースが十分に収集された段階になく、最適な特徴量が何なのかは分かっていないことから、プロトタイプでは既開発の模様特徴を採用した。

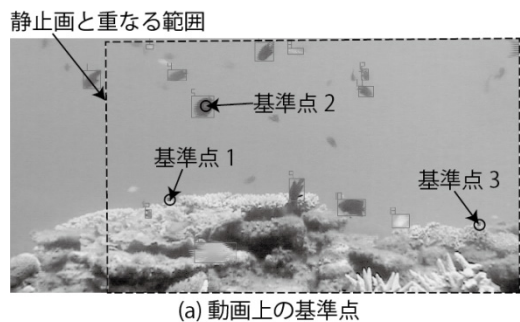
適切な特徴量を見出し、魚影判定の精度向上をはかるには、画像データベースの充実が重要である。そのため、画像データベースを容易に構築できるようにする必要がある。そこで我々は、ラベル付け支援ツールを開発した。当該ソフトウェアでは、魚影対応付けブロックで得られた領域を示す枠を、静止画上に重ねて表示し、魚影の拡大表示、枠の移動・形状変更と、ラベル付けを行うことができる。

## 4. 評価実験

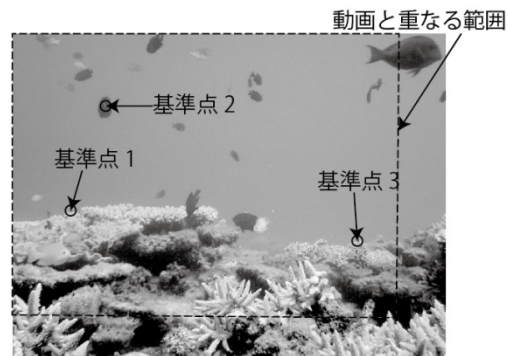
提案手法を評価するため、水中撮影した動画・静止画を用いた評価実験を行った。評価には沖縄県渡嘉敷郡座間味村阿嘉島のマジノハマ沿岸の水深1mに図 4 に示す方法で設置したビデオカメラとスチルカメラにより、2011年8月20日に撮影した26分(14時41分から15時17分)の映像(DVフォーマット)と、同時に撮影した静止画像157枚(10秒毎に撮影、4000×3000画素)を用いた。ビデオカメラとスチルカメラの間はおおよそ20cm離れている。当該地点はサンゴ礁が発達し、魚類が豊富に生息しており、手法の性能評価に好適である。評価実験では、手法を構成する各機能の性能と課題を明らかにすることを目的に、3章で述べたブロックごとに評価した。

### 4.1 魚影抽出ブロックの評価

水中映像では、カメラの設置位置の前方にサンゴがあり、画面の下部3分の1ほどを占めている。動画は DV フォーマット、ワイド化された横 720 画素、縦 360 画素のインターレース画像である。評価対象は、静止画が記録された時刻と同時刻のフレーム画像 157 枚とした。評価の基準となる魚影は、157 枚全て



(a) 動画上の基準点



(b) 静止画上の基準点

図 5: 対応付けの基準点

のフレームを表示し、前後をコマ送りして、目視で見出されたすべての魚影とした。

これに対し、提案手法を用いて魚影抽出を行った結果、正答数 1050 尾に対し、見落とし数 146 尾、間違い数 20 尾となった。よって、適合率 (Precision) =  $1050/(1050+20) = 98.1\%$ 、再現率 (Recall) =  $1050/(1050+146) = 87.8\%$  である。

手法中のパラメータ設定は以下とした。シグマバンドの幅調整係数  $\alpha = \beta = 4.0$ 、平均と標準偏差を求める時間幅 300 秒、魚影の最大速度 20 画素/フレーム、魚影サイズ上限 5000 画素、下限 50 画素、魚影と確定するまでの最小追跡回数 3 である。

見落としは以下の場合に発生した。まず、光のゆらぎの大きいサンゴ表面などでは、シグマバンドが大きくなり、明るい魚の姿を見落とすことがあった。逆に、サンゴの影になる常時暗い場所では、黒い魚を見落とすことがあった。これらは目視でも見出すのが難しく、動画を再生することで始めて見いだせた。また画面上で魚影が重なったり接触したりした場合には、魚影を一つとみなし、見落とすことがあった。

一方で、間違いのほとんどは、大きな一尾の魚を2つ以上に分割してしまう場合であった。光のゆらぎなど背景の明度変化を魚影と間違えることは、ほとんど無かった。

### 4.2 魚影対応付けブロックの評価

#### (1) 動画と静止画の時刻・座標合わせ

動画と静止画の時刻は以下の方法で可能な限り揃えた。まず、計測時間帯の前に撮影された静止画の中で、泳ぐダイバーなど、動きがあって、フレーム単位で変化の見分けがつく画像を選定する。その静止画と同じ動画のフレーム時刻を、静止画の時刻にミリ秒単位で合わせる。動画の各フレーム時刻は、この再設定された時刻から、フレーム数換算で再計算する。

提案手法では、動画と静止画の座標を対応付ける何らかの目印が必要である。今回は、図 5 に示す 3 点を基準点とした。

### (3) 対応付けの精度

魚影対応付けブロックの評価のためには、静止画上の正しい魚影位置が必要である。正しい魚影位置は、3.3 節で述べたラベル付け支援ツールを用いた。評価指標には、以下の領域捕捉率を定義して用いた。

$$\text{領域捕捉率} = \frac{\text{抽出領域と魚影領域の重なる面積}}{\text{魚影領域の面積}}$$

ここで魚影領域とは魚影を囲む最小長方形である。つまり、自動抽出した領域が、正しく魚影を覆っている場合に 100%、全く重なりが無い場合に 0%となる指標である。魚影に対応する全 817 領域に対して、座標変換のみの場合、平均 47.3%、魚影探索を行った場合、平均 70.8%となった。

魚影対応付けブロックのパラメータである、探索範囲係数は  $a = 5.0$ ,  $b = 3.0$  とした。横に広い ( $a > b$ ) とした理由は 2 台のカメラが併置されており、視差ズレが横方向に大きいためである。

評価に用いた母数 817 が、4.1 節で述べた抽出魚影領域 1050 よりも少ない理由は、動画には現れていても、静止画には現れていない魚影がいたためである。そのような領域は、海底や海水など魚影の無い領域に対応付けられる。現れない理由は、スチルカメラでは撮影できない画面の外に魚がいる場合と、2.2 節(1)で述べた時間ズレの影響で、魚影が動画とは異なる位置に移動している場合である。

なお、魚影探索まで行なっても、領域捕捉率が 0%と、対応付けに失敗した領域は 114 あった。これらは主に画面の下部で発生している。これらを分析した結果、原因は以下3つであることがわかった。まず座標変換により、魚影探索の初期位置として与えた座標が、視差ズレの影響により大きく異なっているため、探索範囲に魚影が無い場合である。視差ズレはカメラに近いほど大きい。次に、サンゴなど魚影と区別の付きづらい背景にいる場合、魚ではなくサンゴを選んだ場合があった。これは動画と静止画で色や解像度が異なることも影響している。最後に、2.2 節(1)で述べた時間ズレの影響で、魚の向きや角度が、動画と静止画で大きく異なり、類似画像となっていない場合である。いずれの条件も画面下部で多く発生するため、領域捕捉率が画面下部で低くなっている。

## 4.3 魚種判定ブロックの評価

### (1) 画像データベースの作成

魚影判定ブロックの評価のため、前記の魚影対応付けブロックで得られた領域に対応した、画像データベースを作る必要がある。そこで、魚影対応付けブロックで抽出された 1021 の領域すべてに専門家によるラベル付けを行った。ラベル付けは、ラベル付け支援ツールを用い、前節で述べたように正しい魚影位置に修正した後、行った。なお、抽出領域は魚影だけでなく、サンゴや海底、海中など、魚影以外も含んでいる。なお 4.1 節で抽出した 1050 より少ない理由は、同じ領域に重なった領域を除外したためである。ラベルは「スズメダイ」「その他の魚類」「不明魚」「魚以外」の 4 種類とした。

それぞれの領域数は、スズメダイ 332、その他の魚類 42、不明魚 485、魚以外 162 となった。また、それぞれの領域の平均サイズ(静止画像を  $800 \times 600$  画素とした場合の画素数)は、スズメダイ 2292、その他の魚類 3711、不明魚 980、魚以外 3465 となり、サイズが小さい魚影は判定が困難であった。画面上部ほど不明魚の数が多し。画面上部はカメラから遠い魚影も撮影されるため、判別のできない魚影が多く現れる。なお、魚以外の

領域数が上端に多いのは、静止画撮影時刻が遅れ、魚が画面外に泳ぎ出したためである。

### (2) 魚影とそれ以外の判定性能

魚影判定ブロックの機能として、まずは魚と魚以外を区別する必要がある。そこで自動判定の初歩的なトライアルとして、不明魚も含めた魚影と、魚以外を自動判別することを試みた。自動判定の対象領域は、専門家が魚影位置を修正しラベル付けした領域である。特徴量としては、当所が開発した人物検出で良好な性能を示した模様特徴量[中島 2005]を採用した。具体的には領域画像を縦 64 画素、横 64 画素に縮小し、2 次元 Haar ウェーブレットで得られた LH 成分と HL 成分の計 512 個を用いた。学習器は SVM、カーネルには自然画像に適しているとされる RBF カーネルを採用した。データの 3/4 をランダムに選択して学習に用い、残り 1/4 を評価に用いるクロスバリデーションを、RBF カーネルのパラメータである  $\gamma$  を変えて、各 50 回行った結果を表 1 に示す。魚影とそれ以外の判定は約 94%で可能であることが示された。

今回の撮影ではスズメダイ以外はほとんど撮影できていないため、魚種の判定性能は評価しなかった。魚種毎の判定性能評価や有効な特徴量は今後の課題である。

表 1: 魚影判定精度

$\gamma$	精度 [%]
50	92.8
10	93.8
5	93.7
1	91.8
0.5	90.6
0.1	89.1

## 5. おわりに

本稿では、温暖化影響の把握などを目的とした魚類の生息量調査を支援するために、画像処理を応用した自動定量化手法を新たに提案した。提案手法は Windows 上のソフトウェアとして開発しており、動画・静止画を水中撮影すれば適用可能である。生息量調査の支援に活用することで、調査の効率化に寄与すると考えられる。また適用によって、魚影の画像データベースが充実すれば、魚種の判定まで自動化できる可能性がある。今後は手法の改善とソフトウェアの普及を行なう計画である。

## 参考文献

- [西田 2005] 西田 高志, 中國 明信, 及川 信, 松井 誠一:「近年の海水温上昇による筑前海沿岸魚類相の変化」, 九大農芸誌, 60(2), pp.187-201, (2005)
- [亀井 2009] 亀井 良則, 浜口 昌巳, 菅野 泰久:「岡山県沿岸域で採捕されたなるトビエイの消化管内容物」, 岡山水試報告, 24, pp.32-34, (2009)
- [桑原 2010] 桑原 久実, 明田 定満, 小林 聡, 竹下 彰, 山下 洋, 城戸 勝利:「温暖化による我が国水産生物の分布域の変化予測」, 地球環境, Vol.11, No.1, pp.49-57, (2010)
- [堤 2005] 堤 富士雄, 伊藤 憲彦:「水中または水面における移動体の計数システムおよび計数方法」, 特開 2005-309485, (2005)
- [中島 2005] 中島 慶人:「既設監視カメラを用いた高速な侵入者検知方式」, 電力中央研究所研究報告, R04007, (2005)