

平成23年 5月 10日現在

機関番号： 82401  
 研究種目： 若手研究(B)  
 研究期間： 2009～2010  
 課題番号： 21700213  
 研究課題名（和文）  
 ハイスループット化のための画像処理によるタンパク質結晶成長状態評価手法  
 研究課題名（英文） Protein Crystallization Evaluation Method based on Image Processing  
 for High-throughput Crystallization Experiments  
 研究代表者  
 川端 邦明（KAWABATA KUNIAKI）  
 独立行政法人理化学研究所・分子情報生命科学特別研究ユニット・専任研究員  
 研究者番号： 90301754

## 研究成果の概要（和文）：

本研究では、タンパク質結晶化のハイスループット化を支援するための画像処理を用いた結晶成長状態評価の自動化に関わる手法の開発およびそれらの手法を統合したシステム・ソフトウェア化を行った。本研究により、自動観察システムTERAが14,000枚の画像を一日で取得するのに対して、17,000枚/日超程度の処理性能を持ちつつ、成長状態の5段階への分類が行われる結晶成長評価支援の効率化に貢献可能な技術が開発された。

## 研究成果の概要（英文）：

In this research subject, protein crystallization evaluation method based on image processing for high-throughput crystallization experiments. Developed image processing methods for automated evaluation are integrated as an integrated system software. As the result, the system software realizes five group categorization and shows the processing performance (17,000 pictures/day) over the one (14,000 pictures/day) of automated observation system TERA, which was developed by RIKEN.

## 交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2009年度	1,800,000	540,000	2,340,000
2010年度	1,600,000	480,000	2,080,000
総計	3,400,000	1,020,000	4,420,000

## 研究分野：総合領域

科研費の分科・細目：情報学・知覚情報処理・知能ロボティクス

キーワード：画像処理，パターン認識，タンパク質結晶化実験，結晶成長評価，システム・ソフトウェア，ハイスループット

## 1. 研究開始当初の背景

タンパク質結晶構造は、生物学，分子生物学等

の先端科学研究分野に影響を及ぼす重要な知見である。タンパク質の立体構造を解析する方法の一つとして、X線結晶構造解析法がある、この手法では小さな分子量のタンパク質単体から、超巨大複合体や蛋白質複合体までの解析が可能である。X線構造解析法には、良質のタンパク質結晶が必要である。しかしながら、結晶化条件が未知であることから、結晶化作業は網羅的に探索が行われることが多く、構造解析のボトルネックとなっており、結晶化作業のハイスループット化が求められてきた。このことから、近年、結晶化作業の自動化を試みたシステムの開発が進められており、その典型例に理化学研究所播磨研究所において開発された自動結晶化観察システム:TERAがある。しかしながら、TERAにより仕込まれた結晶化サンプルの成長状態評価は目視で行われているのが現状である。TERAは14,000条件/日もの画像撮像能力をもつことから、ハイスループット化には、結晶化サンプルの成長状態評価の自動化が必要不可欠となっていた。

## 2. 研究の目的

TERAで行われる結晶化実験において結晶化サンプル内にX線回折実験対象となりうる結晶が析出したか等の評価は、専門家によって目視で行われている。TERAは、14,000条件/日もの画像を撮像する能力をもつことから、効率的な構造解析を実験するためには、結晶化サンプルの成長状態評価自動化が必要不可欠でとなっている。

そこで本研究では、タンパク質結晶成長状態評価作業を支援するための画像処理に基づいた一連の処理を自動化・統合化したシステム・ソフトウェアの開発を行うことを目的

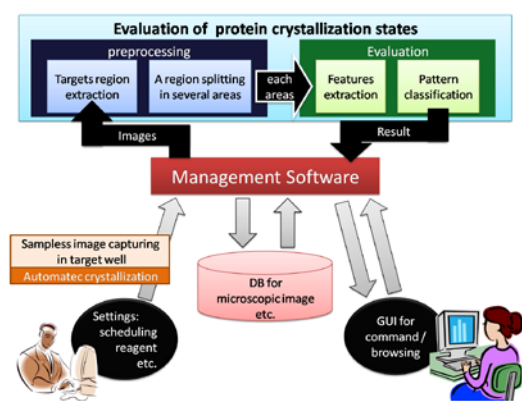


Fig System configuration

of protein crystallization states evaluation system

とした。

特に本研究では、TERAにより撮像された結晶化サンプル画像に対して、自動で結晶成長状態評価を行い、評価結果を視覚的に表示

することで評価作業を支援するソフトウェアを構築したので、報告する。

## 3. 研究の方法

実際に行われているタンパク質結晶化作業、結晶成長状態評価を参考にすると、自動評価システムとして最低限満たすべき機能要件は以下のようになる。

- 1) 設定した条件でのタンパク質自動結晶化
- 2) カメラによるウェル範囲内自動撮像
- 3) 撮像された画像から評価領域の抽出処理
- 4) 結晶成長状態の自動評価処理
- 5) 評価結果の保存処理
- 6) 評価結果の検索・閲覧

ここで本研究では、1)~2)についてはTERAを用いることを前提とし、3)~6)の機能の達成を目指す。

ここで前述の議論に基づいて我々は、前項の要件を満たすタンパク質結晶成長状態評価システムの設計を行った。このシステムは大きく自動撮像部、自動評価部、データベース部、閲覧用GUI部から構成される。各部の機能を述べる。

顕微鏡撮像部は、顕微鏡によりプレート内の各ウェルの画像を取得するものであり、本システムでは、画像はTERAの撮像系により取得することを想定している。撮像される画像は、プレート内の各ウェルを40倍に拡大して撮像した1392[pixel]×1040[pixel]のJPG形式画像である。

結晶状態評価部は、指定されたフォルダの監視を行い、新たな画像ファイルが追加されたのを検知すると、その画像を取得し結晶状態の評価を行う。

ここでカメラの位置やフォーカスの精度により、画像中のウェル底面は必ずしも同じ位置、大きさではないことから、評価の前に画像処理によって評価領域を抽出する必要がある。評価領域は、円領域内の彩度分散値が最小となる半径および中心座標を導出することで抽出される。また、評価領域抽出の速度向上を図るため、縮小した画像を用いて抽出を行うこととした。

基礎実験の結果より、1369枚の画像からウェル位置抽出成功と判断した画像1322枚のうち、1320枚が正しく抽出できることが分かった。

つづいて結晶化に関する評価は、抽出した評価領域を参考に、さらに13の小領域に分割し、それぞれの小領域で独立して評価を行うこととした。

評価には我々が開発を行ってきた手法を応用し、初期成長状態を主とする5段階(A:クリア, B:沈澱(i), C:沈澱(ii), D:沈澱(iii), E:非晶質の粒+微結晶+結晶,)の評価を行う。各小領域に対してテクスチャ解析を用いて抽出した28種類の特徴量を用い、線形判別分析を用いた階層構造からなる多クラス識別機を構築した。これまでの基礎実験において、専門家の評価に基づいた874枚の画像に対する5段階評価の性能が平均90%程度あることが確認されている。

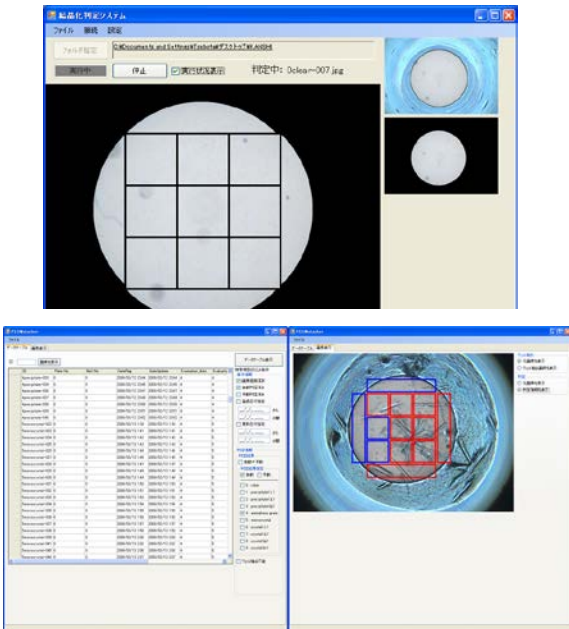


Fig. Function of search and browsing images by GUI

データベース部では、各サンプルの画像、プレート番号・ウェル番号、各小領域の結晶成長状態評価結果などの他に、抽出した評価領域の円半径・中心座標等の処理結果をデータベースに登録する。本システムではデータベースソフトウェアとしてSQL Serverを用いた。

撮像されたサンプルの、画像および結晶状態評価結果はデータ閲覧用GUI部によって確認が可能となっている。データ閲覧用GUI部では、データベースにアクセスし一覧を取得し、一覧から選択された画像を表示する。特定の評価結果や、データベース登録日時などの条件で絞り込んで検索も行え、任意のサンプル画像を簡単に選択可能となっている。また、選択したサンプル画像に対して評価領域の抽出、小領域ごとの評価結果を色つきの枠で表示などのマスクをかけることができ、結晶状態評価結果を視覚的に閲覧することが可能となっている。

#### 4. 研究成果

試作したシステム・ソフトウェアを用いて、結晶化状態評価からデータベース登録までの一連の流れの所要時間を計測する性能評価実験を行った。

試作ソフトウェアはIntel® Core™ i7 965 @ 3.20GHz, 8GB RAMのPC上に実装されている。



この試作システム上に用意した1369枚の画像に対して、評価、データ登録を終えるまでシステムを動作させる。実験は、評価領域抽出の際に用いる画像サイズを変化させて、各サイズでの所要時間を計測した。ここで所要時間は、最初の画像が登録された時刻から、最後の画像がデータベースへ登録された時刻までとした。

Fig. Overview of prototype system

基礎実験の性能の結果から、1/16サイズの画像を用いた場合に、抽出処理エラーが最も少なくなったため、評価領域抽出処理に用いた画像サイズを1/16とした。また、抽出した画像に対する5段階に対する評価処理の性能は平均82%程度の正答率であった。

一方、一連の処理として試作ソフトウェアにより実験を行った結果、17,292枚/日(0.200枚/sec, 5.001sec/枚)の性能があることが分かった。また所要時間0.21分/枚の処理であり、6,857枚/日の画像の評価が可能である。TERAの画像収集能力14,000枚/日であることから、現在一般的であるPC程度の計算機性能でも十分な処理能力があることがわかった。

このことから本研究により、自動観察装置TERAで撮像された画像を対象とした、結晶成長状態自動評価のためのシステム・ソフトウェアの開発が行われた。各課題に対する画像処理手法や情報処理手法を統合した自動評価処理に加え、結果を視覚的に確認可能なGUIを導入することにより、結晶成長状態評価作業の支援に寄与する性能を示すことができた。

今後の課題は、今回の実験で誤評価した小領域に関する分析に基づいた手法の改善や他の評価手法との統合による性能の評価などがあげられる。また、それに付随して現在の処理性能を維持することも併せて考慮する必要がある。

#### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 1件)

① K. Kawabata : "Automatic detection of crystallization evaluation region by using image processing", *Sensor Review*, 査読有り, 2011(accepted)

[学会発表] (計 2件)

① 川端邦明 : "タンパク質結晶成長状態自動評価ソフトウェアの開発", 計測自動制御学会システムインテグレーション部門講演会講演論文集(SI2009), 730-733, 2009年12月25日, 東京

② 川端邦明 : "タンパク質結晶成長状態評価システムの開発—自動評価処理のためのソフトウェア構築—", ロボティクス・メカトロニクス講演会'09講演論文集(CD-ROM), 2A1-B05, 2009年5

月 26 日, 福岡

6. 研究組織

(1) 研究代表者

川端 邦明 (KAWABATA KUNIAKI)  
独立行政法人理化学研究所・分子情報生命科学特別研究ユニット・専任研究員  
90301754

(2) 研究分担者

なし

(3) 連携研究者

なし