2章 センサ・撮像信号処理技術

伊賀 希一郎[†], 清 正 史[†], 加治木 圭太[†], 堤 則 彦[†]

キーワード：画像処理, 撮像センサ, HDR, LTM, ダイナミックレンジ, ノイズ除去

1. まえがき

デジタルカメラは、嗜好性と実用性の2面性を持っている。家庭や趣味で使われる一般的なデジタルカメラのような嗜好性の高いところで必要とされる撮影のための画像処理技術と実用性いわゆる監視カメラなどの用途に必要な画像処理技術には、一定の差があることは一般的である。画像処理に関連する事業を広い市場への適用が必要とされる画像処理事業を行うためには、上述の両面に対していかに早期簡潔に課題を解決し、直面する市場に対し最適な画像を提案することができるかがカギとなる。そして、先進的な画像処理技術は、そのために必要となる量産製造ができる環境をいかに提供することができるかが大きな市場優位性であり、事業における継続した価値の連続性を保つ重要なポイントとなる。

ソシオネクストミルビュー事業部では、汎用画像処理SoCソリューションミルビュー (Milbeaut) を開発し販売している。このミルビューシリーズは、デジタル一眼レフから携帯電話カメラ、アクションカメラそして監視用途まで幅広い分野を市場ターゲットとしており、画像センサ入力からイーサネットへの配信まで、カメラ画像処理商品に必要な技術をすべて実装している。搭載する技術も異なる市場ターゲットを超えて多彩に活用することで早期提供を実現し、お客様の生活を充実させることによる社会貢献に努めている。

本稿では、監視カメラ向けに重要となる機能に着目し、撮像から蓄積・配信までの中で重要な視認性を上げる部分に重点を置いた課題を解決する技術について述べる。特に、撮影物体への追従性を上げるための高速センサへの対応・既存の市場センサでは、視認性限界のあるダイナミックレンジへの対応そして解像度とコントラストをいかに保持し続けるかの対応を中心に、基本的な画像処理とそれに伴い

生じる課題の解決手段について紹介する。

2. 撮像のための画像処理技術と課題

2.1 CMOSセンサのローリングシャッタ歪みに対する対策

通常のCMOSセンサにおける電子シャッタの動きは、センサ全面で一斉にシャッタが切られるわけではなく、センサの左上端の画素からラスト順にシャッタが切られる、いわゆるローリングシャッタの動作をする。メカシャッタが付いていれば、メカシャッタを切って撮像素子への入射光をセンサ全面で一斉に遮断すれば問題ないが、メカシャッタがない場合は、電子シャッタに頼らざるを得ない。またメカシャッタがある場合においても動画撮影時には、撮影の間メカシャッタを使い続けることはできず、電子シャッタに頼ることになる。

電子シャッタはローリングシャッタ動作をするため、画面の上のほうが早いタイミングでシャッタが閉じられて、画面の下の方は遅く閉じられる。静止している被写体ならば問題ないが、動きのある被写体の場合には、この時差が撮像画像の歪み (ローリングシャッタ歪み) を生むことになる。

電子シャッタのこの時差が1/30 sの場合は、33 msの間の被写体の動きが、ローリングシャッタ歪みとして現れる。このローリングシャッタ歪みを、メカシャッタを使用せずに発生させないようにするためには、電子シャッタの時差を短縮すればよい。それに応じて歪みも減らすことができる。したがって、CMOSセンサからの取り込み速度を究極に高速化する方法が歪み解消のための課題となる。

もしCMOSセンサが1 Gpix/sで取り込むことができると仮定すると、4K画像 (画素数約8 Mpix) の場合には、8 Mpix/(1 Gpix/s) = 8 msで画像一面を取り込める。ローリングシャッタ歪みは約1/4に抑えられることとなる。

このことをCMOSセンサとISP (Image Signal Processor) で実現することを考える。まずCMOSセンサで電子シャッタを切ると、ラスト順に画素値データがセンサからISPに流れてくる。

しかし、1 Gpix/sでISPに流れ込んでくる画像データを

* 本稿の著作権は著者に帰属致します。

[†] 株式会社ソシオネクスト ミルビュー事業部 開発部
"Imaging Technology for Capture and Processing" by Kiichiro Iga, Masafumi Sei, Keita Kajiki and Norihiko Tsutsumi (Milbeaut Business Unit Development Department, Socionext Inc., Yokohama)

図1 センサとISPの画像データフロー

入力と同一速度でISPが画像処理をすることができれば良いが、画像データが流れてくる速度が高速であるため、ISPの処理能力不足で同一速度で画像処理できず、画像データを捌ききることができない場合がある(図1(a))。

そこで、現状では次の方法が一般的である。ISPに流れ込んでくる画像データをISPの入り口ですぐにDRAMに出力して、そこに一旦蓄えるようにする。その後、ISPが捌ける速度でDRAMの画像データを再入力して画像処理を行うようにする(図1(b))。フレームレート1/30fpsの場合、ISPの処理速度は1フレーム当たり33msまで抑えることができるようになる。

2.2 高感度撮影時の長波長ノイズに対する対策

高感度撮影時には、低感度では目立たなかった空間的に長波長で現れるノイズが気になってくる。

ノイズ成分を扱う場合、画像を周波数帯域別に分解して、それぞれの周波数帯域に対して別々にノイズ抑制のパラメータ調整できるようにする。各周波数帯域では、

- (1) 画像の凹凸がエッジであるかどうか推測する輪郭度
- (2) 画像の明るさ

などの値に応じて、ノイズ抑制強度を調整できるようにする。最後に周波数帯域別の情報を統合して、ノイズ抑制を実際に行う(図2)。このようにして、周波数、輪郭度、明るさの3次元的にノイズ抑制をコントロールできるようになる(図3、図4)。

2.3 エッジ強調時のドット抑制対策

通常、エッジ強調は、ラプラシアンフィルタ等のHPFで画像の高周波成分を抽出して、この高周波成分にスケール調整などの加工をしてエッジ強調付加成分を作り、元画像に足し込むことで行われる(図5(a))。

この時、エッジ強調後の画像を見ると、エッジは強調されているが、同時に不自然なドット状のアーティファクトが目につくことがある。この原因は、HPFでエッジ強調前画像の2階微分係数を抽出して、それを強調するようにエッ

図2 ノイズ抑制のデータフロー

図3 ノイズ抑制OFF

図4 ノイズ抑制ON

図5 エッジ強調時のドット抑制対策

ジ強調処理がされて、エッジ強調前画像では目につかなかった小さな凹凸が目に見えるようになったことにある。

デジタルカメラの画像処理では、解像度を上げるような努力を強いられるため、RAW画像からエッジ強調前画像を作成するまでの間において、2階微分係数まで滑らかにする(エッジを鈍らせる)ような画像処理(ノイズ抑制、デモザイク等)することが困難になっている。そのため、不自然な2階微分係数信号が残り、これがドット状のアーティファクトとして画像に表れるようになる。

これに対処するために、次のようなドット抑制処理を追加する方法も存在する(図5(b))。

高周波成分信号を見て孤立点(手段の一例として、注目画素の周囲8点の画素と比べて、注目画素が大きい、または小さい点)を見つけた場合、その孤立の程度を抑制するような値に置き換える。

このような方法で、ドット状のアーティファクトを抑制することができる。方法はこれらに限らないが、何らかの手段で対策を入れることが望ましい。

3. 視認性向上のための画像処理技術

本節では、ダイナミックレンジ拡大による視認性向上技術について紹介する。

3.1 市場要求と合成指針

自然界のダイナミックレンジは、カメラの1回の露出によって得られるダイナミックレンジに比べてはるかに広い。ため、カメラによって得られる映像に白飛びや黒つぶれが含まれてしまうことがしばしば発生する。その解決策として、一般的には異なる露出条件で撮影された複数枚の画像を合成することで自然界のダイナミックレンジを実現する手段がとられるが、その合成画像は市場により微妙に異なるニーズを持っている。デジタルスチルカメラ(Digital Still Camera, 以下DSCと省略)市場においては、より自然な諧調を持った画像が特に要求され、近年のテレビ市場

図6 照度の代表例

では、すべての諧調に情報を持った多ビット画像が要求され始めているなど、市場によって複数のニーズがあるが、監視カメラ市場では特に、明部暗部のオブジェクトが1枚の画像の中でより鮮明に見えることを要求される。

監視カメラ市場では、ビルのエントランスや地下駐車場など自然界の中でもとりわけダイナミックレンジの大きいシーンの元で撮像される機会が多い。このようなシーンで良好な解像を得るためには照度比約1,000~10,000倍の端と端に位置するオブジェクトを1枚の画像へ同居させる必要がある(図6)。

次に、短露光と長露光で撮影した2枚の画像を合成する場合について説明する。すべての諧調へ情報を持たせる際には、図7(a)に示されるように、片方の画像(図では短露光)の入力レベルへ合わせこみ、良好な解像を得ている画素を利用する手法が一般的である。しかしこの手法では、諧調の連続性は良好なもののダイナミックレンジが広くなるにつれて、利用できるレンジが狭くなり、強力なWDR(Wide Dynamic Range)処理を後段に構える必要がある(図8(a))。一方、明部と暗部の再現性に着目した際には、図7(b)に示されるように点線の区間をうまくつなぎ合わせることで、良好な解像を持った領域を同居させることができる(図8(a))。諧調の連続性という観点では図7(a)

図7 画像合成時の入出力例

図8(a) 図7(a)の手法で合成した画像

図8(b) 図7(b)の手法で合成した画像

に劣るものの、比較的小さいハードウェアコストで実現できることが利点である。

3.2 市場間で共通の技術

複数枚の画像を合成する際には、原理的な課題が2点ある。一つはフレームの位置ずれ（以下、GM: Global Motionと表現）、もう一つはフレーム間の被写体動き（以下、LM: Local Motionと表現）である。

GMの検出および補正で考慮すべき主なポイントとしては、位置ずれ発生源とセンサ・レンズの特性が挙げられる。監視カメラにおける主な位置ずれ発生源は、カメラの旋回や振動、DSCにおける主な位置ずれ発生源は撮影者の手ブレである。発生源に適した調整範囲で最適化を図る一方、ローリングシャッタ歪みやレンズ歪みの影響を考慮し、最適な補正用ハードウェア構成を選択する必要がある。

LMの検出および補正で考慮すべき主なポイントとしては、フレームレートと解像度である。フレームレートが低くなるほどフレーム間の時間が開き動きは大きくなるが、昨今のカメラシステムではより高いフレームレートに移行しているため、懸念は小さくなっていくだろう。一方、解像度はより上がる傾向にあるため、より多くの画素の補正が必要になる。一般的なLMの補正手法として、動き補償（Motion Compensation）が挙げられる。単位領域毎の動き

図9 ピラミッド型検出の例

図10 GTMの例

ベクトルを見積もり、その結果から動きを補正するものだが、解像度の向上は単位領域で捕捉できる動きが小さくなることを意味する。この課題の解決には縮小画像を用いたピラミッド型の動き検出が有効である。縮小スケールに応じて捕捉する動きの範囲を柔軟に制御することが可能である（図9）。

3.3 トーンマッピング処理

デジタルカメラや監視カメラのコントラストやダイナミックレンジを調整し視認性を向上させる技術として、トーンマッピング技術が一般的である。その代表としてグローバルトーンマッピング（以降GTMと表現）とローカルトーンマッピング（以降LTMと表現）が広く知られる。

GTMは、画像全体に一律の明るさ調整を施す手法（図10）で、トーンマッピング調整テーブルを一意に定義して画素値を入力とした1対1の出力を演算する。なお、調整テーブルは、可変パラメータとして自由な曲線を描けることが望ましい。

ハードウェア実装が容易でリアルタイム性に優れる反面、画像の凹凸を考慮しない一律処理となるため、ローカル領域に着目した場合にコントラストの低下を招きやすいデメリットが生じる。

LTMは、注目画素近傍の明るさを判定し、画素毎にコントラスト調整を施す手法（図11）で、注目画素の画素値のみに依存せず局所的な明るさを考慮した出力を演算する。

画像の凹凸を保存（強調）したコントラスト調整が可能である反面、演算時の参照画素範囲が広がるため、ハードウェア実装が複雑に（ハードウェアコストが高く）なりやす

図11 LTMの例

図13 LTM処理フローの例(2)

図12 LTM処理フローの例(1)

図14 LTM処理フローの例(3)

いデメリットが生じる。しかし、撮像センサの高画素化や光学系の低価格化に伴うレンズのMTF低下等により、画像処理による高画質化が求められており、LTMの重要性が増している。そのため、デジタルカメラや監視カメラの動画の視認性向上のため、画質優位性のあるLTMを活用するケースが増えており、その一例として動画にLTMを適用するための施策を紹介する。

LTMでは、注目画素とその近傍画素の画素平均値を演算することで、注目画素が存在する領域の照明光の明るさを判定する。近傍画素の参照範囲が狭いと明るさを正しく判定できないため、より広範囲の近傍画素を参照する必要がある。ISPで画像処理する際、広範囲の画素を参照するためには、参照画素範囲分のディレイバッファを用意する手法や、画像を縮小し空間フィルタを適用して拡大することにより模擬的に参照範囲を広げる手法が考えられる。後者は前者よりもハードウェア実装コストを抑えることが可能なため、後者を想定した手法(図12)を説明する。

縮小画像を拡大することで、注目画素近傍の明るさを判定するが、縮小と拡大により被写体のエッジが潰れてしまうため、エッジ近傍の明るさ判定を間違えてしまうアーティファクトが発生する。これを解決するため、注目画素の輝度値を参考にして、低周波エッジをある程度復元させながら拡大画像(照明光)を作るのが望ましい(図13)。

図13のような画像処理システムを構築するためには、フル解像度画像と縮小画像をSDRAM等のメモリーに一旦格納する必要がある。動画サイズは近年4Kサイズ(例: 3,840 × 2,160)や8Kサイズ(例: 7,680 × 4,320)に拡張され、安価なハードウェアシステムのSDRAMバンド幅では限られた時間内に処理し切れないケースが発生する。そこで、

SDRAMにフル解像度画像を格納せず、縮小画像のみSDRAMや内部RAMに格納する手法(図14)が考えられる。

演算対象となる縮小画像は1フレーム前に生成したものとなるため、手ブレや被写体ブレによりフル解像度画像と縮小画像との間で被写体の位置関係が異なってしまう(位置ズレ)問題が生じる。そのため、画像を縮小拡大する際、位置ズレを緩和する手法を取り入れるのが望ましい。

なお、本システムを導入することにより、前段処理～LTM～後段処理を一気通貫で処理することが可能となり、画像表示もしくは画像保存までのレイテンシが短縮される効果が得られる。

4. むすび

本稿では、撮像の画像処理としてノイズ抑制、エッジ強調、画像合成、トーンマッピング技術を紹介した。画像センサの進化により、SN比改善とダイナミックレンジ拡大が進んでいるが、動画の高画素化によるノイズ増加や監視カメラ等での視認性向上の観点で画像処理技術の更なる進化が必要と考える。特に視認性向上については、出力媒体のビット深度が従来の8ビットから10、12ビットへ拡張される傾向があるため、ユーザがダイナミックレンジ拡大を滑らかな諧調で感じ取れるようになりつつある。さらに、近年、単眼画像センサでの複数露光画像取得や、複眼画像センサによるダイナミックレンジ拡大技術等も報告されており、技術革新がより一層進んでいくと考えられる。撮像の画像処理において連続的な進化と価値創造により技術革新を推し進めていきたい。

(2015年11月9日受付)

図15 LTM OFF

図16 LTM ON

伊賀希一郎 いがき いちろう 1991年、富士通VLSI(株)入社。汎用DSPコア開発に従事後、1998年より、画像処理デバイスの開発を開始。2006年、富士通(株)先端ソリューション事業部に所属。以降、画像処理アーキテクチャとR&Dを担当。現在、(株)ソシオネクストミルビュー事業部開発部長。

清 正史 きよまさ 2007年、富士通(株)電子デバイス事業本部へ入社。カメラ画像処理アルゴリズムの開発に従事。現在、(株)ソシオネクストミルビュー事業部開発部に所属。

加治木圭太 かじき けいた 2004年、富士通LSIソリューション(株)入社。PC向け地上デジタル放送コンテンツ保護技術開発に従事後、2010年より、カメラ画像処理アルゴリズムの開発およびファームウェア開発に従事。現在、(株)ソシオネクストミルビュー事業部開発部に所属。

堤 則彦 つつみ のりひこ 2007年、富士通(株)電子デバイス事業本部入社。画像処理デバイスの論理設計業務に従事後、2009年より、画像処理アルゴリズムの開発を担当。現在、(株)ソシオネクストミルビュー事業部開発部に所属。