# 光応用技術研修会

「デジタルカメラの画像処理」

2017年6月5日 蚊野 浩

# 目次(1/2)

- 1.1 デジタルカメラの構成
- 1.2 写真の撮影から表示・保存
- 1.3 画像の標本化と量子化
- 1.4 画像データの表現
- 1.5 空間領域と周波数領域の画像 2.5 しぼりと画像
- 1.6 線形空間フィルタ
- 1.7 空間フィルタと周波数フィルタ
- 1.8 点拡がり関数(PSF)
- 1.9 線形フィルタの重ね合わせ
- 1.10 ぼけ画像の復元
- 1. 11 MTF(変調伝達関数)
- 1.12 画像の幾何変換と補間

#### 1章 デジカメ画像処理の基礎 2章 デジカメの基本機能と画像

- 2.1 デジカメ画像処理の全体像
- 2. 2 位相差AF
- 2.3 コントラストAF
- 2. 4 ホワイトバランス補正
- 2.6 シャッターと画像
  - 2.7 手ぶれ補正と画像

# 目次(2/2)

- 3.1 カラーフィルタアレイとデモザイク
- 3.2 画像に発生するノイズ
- 3.3 ノイズ除去
- 3.4 PSNRによる画質評価
- 3.5 シェーディング補正
- 3.6 歪みの補正
- 3.7 ガンマ補正
- 3.8 トーンマネジメント
- 3.9 高ダイナミックレンジ(HDR)画像
- 3.10 顔画像の検出
- 3.11 さまざまな顔画像処理
- 3.12 移動物体の追跡
- 3. 13 超解像処理
- 3.14 パノラマ画像

#### 3章 デジカメ画像処理の各論 4章 画像映像信号の圧縮技術

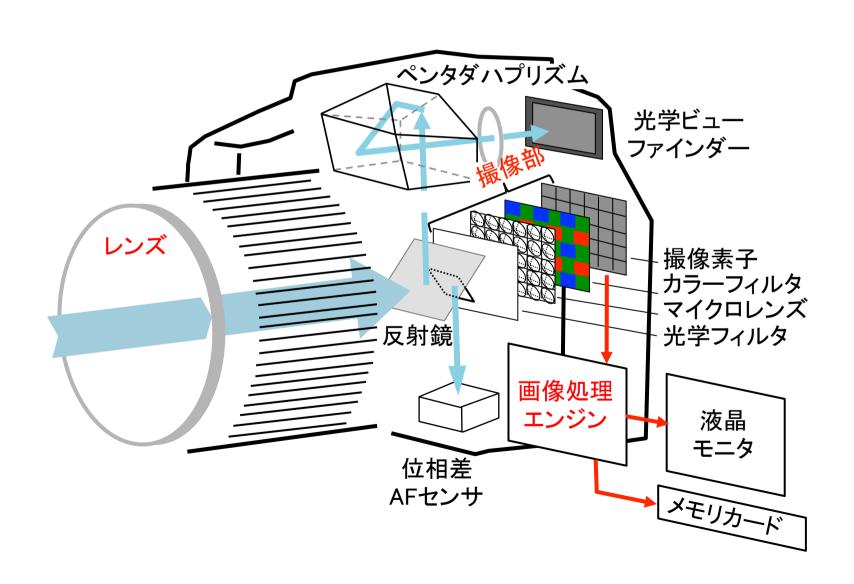
- 4. 1 JPEGによる静止画の圧縮
- 4.2 JPEG符号化器の構成
- 4.3 2次元DCTの計算手順
- 4.4 2次元DCTの効果
- 4.5 動画像
- 4.6 映像(動画)圧縮技術
- 4.7 MPEG符号化のデータ構造
- 4.8 動き予測と動き補償
- 4.9 MPEG符号化の処理フロー
- 参考になる書籍

# 1章 デジカメ画像処理の基礎

# 1. 1 デジタルカメラの構成

- デジタルカメラはレンズ、撮像部、画像処理エンジンを主要部品とする装置である。撮影した画像・映像を液晶モニタに表示して確認したり、そのデータをメモリーカードに保存することができる。
- その他の構成部品はカメラの種類(レンズ交換型やレンズー体型)によって異なる。一眼レフは、反射鏡を介して光学ビューファインダ(OVF: Optical View Finder)によって被写体を直視できることと、位相差AF(Auto Focus)センサを有することが特徴である。撮影時には、反射鏡が跳ね上がり、光像が撮像部に導かれる。
- 一眼レフ以外のカメラでは、撮影前の被写体確認を液晶モニタで行うものが多く、これをライブビューと呼ぶ。ファインダで確認する場合には電子ビューファインダ(EVF: Electric View Finder)を用いる。AFはコントラスト方式を用いるものが多い。

# 1. 1 デジタルカメラの構成



# 1.2 写真の撮影から表示・保存

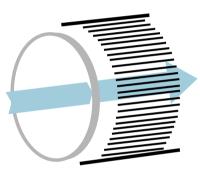
- デジタルカメラにおける最終画像生成までのプロセスを概観する。
  - 被写体(情景)から明暗(光の強度)と、絵柄の変化(空間周波数)が広い光線がレンズに達する。
  - レンズの働きによって光線が結像する。このとき、ぼけや歪み、不自然な陰 影(シェーディング)などが加わった光の像が撮像部の表面に形成される。
  - 撮像部は、光像に帯域制限や色分解・標本化・量子化などの処理を施し、 生のデジタル画像を生成する。生のデジタル画像は、光像自身に存在した 不具合や、デジタル化におけるノイズの混入などの影響で劣化している。
  - 生のデジタル画像を観察に適した最終画像に変換する処理が、デジタルカメラの画像処理である。この中には「生画像の現像」注)と呼ばれる処理も含まれる。これらの処理は「画像処理エンジン」と呼ばれる部品で実行される。画像処理エンジンはCPU・DSP・専用回路・ROM・RAMが混在したシステムLSI(SoC: System on Chip)である。
  - 最終画像は、付属の液晶モニタや外部ディスプレイで観察する。また、撮影 データをメモリカードに保存する。この時、画像・映像圧縮技術を用いて、 データ量を削減する。

# 1.2 写真の撮影から表示・保存

情景:明暗と解像度の範囲 が非常に広い発光体



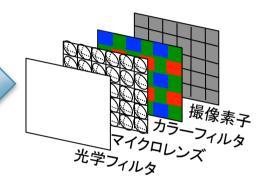
レンズ: 原シーンを光像に 変換し、ぼけ・歪み・シェー ディングなどが加わる

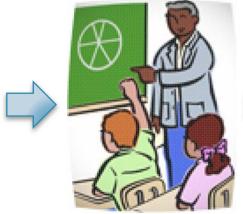


劣化した光像



撮像部:光像に対して、帯域制限・色分解・量子化・標本化を加え、デジタル画像に変換する





生のデジタル画像



デジタル画像処理:生の デジタル画像を、見やす い画像に変換する



デジタルカメラの最終画像



液晶 モニタ



メモリカード

画像・映像信号の圧縮

### 1.3 画像の標本化と量子化

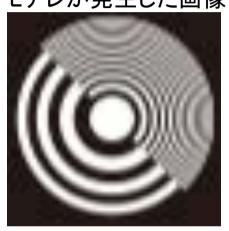
- 撮像部で行われる、CCDやCMOSセンサによる画像の標本化と量子 化は特に重要である。
- 標本化は、アナログ画像である光像を縦方向と横方向で空間的に離散化する処理である。標本化定理により、アナログ画像に含まれる最高周波数の2倍以上の周波数でサンプリングすれば、アナログ画像と標本化された画像は等価である。
- 標本化定理を満たさない場合、忠実に再現することが不可能な高周波成分は、デジタル画像にモアレ(エイリアス、折り返し歪み)と呼ばれる副作用を生じる。これを軽減するために、撮像部に光学ローパスフィルタを配置することがある。
- 標本化に続いて、アナログの信号値をデジタル値に変換する量子化が行われる。量子化のビット数が12ビットであれば、2<sup>12</sup>=4096段階の明暗を表現することができる。ビット数が少ない場合、階調がなだらかに変化する部分に境界線が現れやすい。これを<mark>疑似輪郭と呼ぶ</mark>。

# 1.3 画像の標本化と量子化

モアレが無い画像



モアレが発生した画像



疑似輪郭が無い画像





疑似輪郭が発生した画像

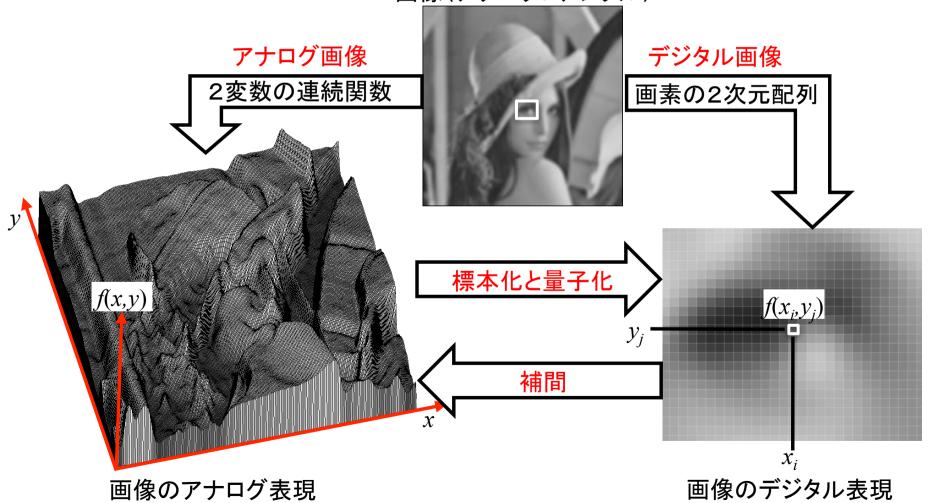


# 1.4 画像データの表現

- 写真画像は、レンズが形成する光像から、撮像部で生成されるデジタル画像、画像処理エンジンが出力する最終画像へと、段階的に処理される。このとき、レンズが作る光像はアナログ画像、撮像部の出力以降はデジタル画像として表現・処理される。
- アナログ画像は、連続的な水平・垂直位置(x,y)に対して、光の明るさに対応する連続量f(x,y)を与える連続領域での2変数関数として表現できる。
- デジタル画像は、離散的な水平・垂直位置 $(x_i,y_j)$ に対して、量子化された 値 $f(x_i,y_j)$ を与える離散領域での2変数関数として表現できる。同じことであるが、画素値の2次元配列として表現されることもある。
- アナログ画像からデジタル画像への変換では、水平・垂直位置の標本化とデータ値の量子化が行われる。デジタル画像からアナログ画像への変換にはデータの補間処理が必要である。これらが標本化定理の条件に基づいて実行されれば、アナログ画像とデジタル画像は等価である。(厳密に言えば、量子化された時点で標本化定理の条件を満たしていない)

# 1. 4 画像データの表現

画像(アナログ/デジタル)



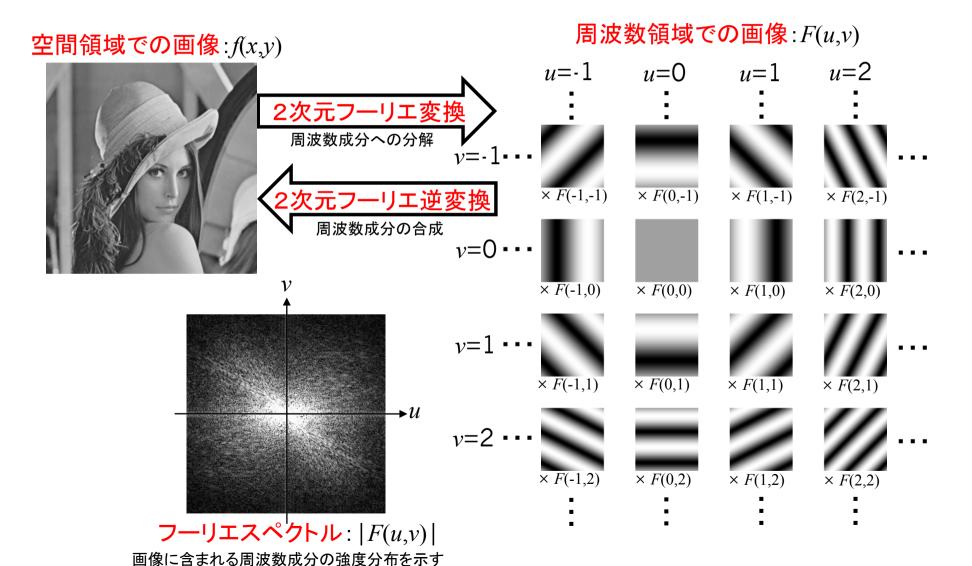
連続領域での2変数関数 f(x,y)

離散領域での2変数関数  $f(x_i, y_i)$ 

#### 1.5 空間領域と周波数領域の画像

- 通常、画像は前ページのように表現される。この表現を、次に述べる周波数領域の画像と対比させて、空間領域における画像と呼ぶことがある。
- 空間領域の画像を、フーリエ変換によって周波数領域に変換できる。変換結果 を 問波数領域における画像と呼ぶ。画像をフーリエ変換することは、画像を2次 元周波数の基底画像に分解し、それぞれの強度と位相を求めることである。
- 空間領域の画像と周波数領域の画像は2次元フーリエ変換と2次元フーリエ逆変換によって相互に変換できる。デジタル画像に対するフーリエ変換は離散フーリエ変換と呼ばれる。離散フーリエ変換には高速フーリエ変換と呼ばれる高速アルゴリズムが存在する。
- 画像のフーリエスペクトルは、周波数領域の画像に対して、水平・垂直周波数とその2次元周波数の強度の関係を表したものである。
- 画像を周波数領域に変換することで、空間領域での処理に比べて、処理の高速化が可能な場合がある。また画像の周波数特性を容易に処理・理解できる。

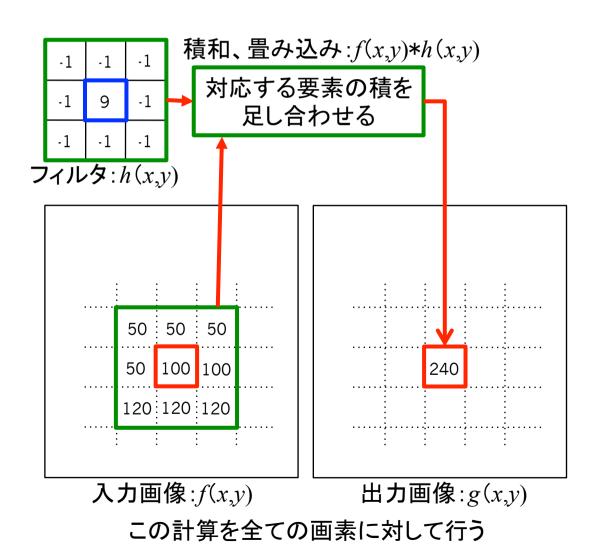
# 1.5 空間領域と周波数領域の画像



## 1. 6 線形空間フィルタ

- デジカメで利用される、最も基本的な画像処理は線形空間フィルタである。これは入力画像にフィルタと呼ばれる2次元係数を作用させる。
- 線形空間フィルタは、注目画素の周囲にフィルタと同じ大きさの近傍領域を設定する。近傍領域の画素値と対応するフィルタ係数の積を計算し、全ての積を足し合わせる。その計算結果を注目画素の新しい値とする。この処理を、入力画像の全画素に対して実行する。
- 上記の計算を積和演算、あるいは畳み込み演算とも呼ぶ。
- 線形空間フィルタの効果・特性は、フィルタ係数によって決まる。画像にぼけを与えるフィルタは平滑化フィルタである。平滑化フィルタの中で、ガウス分布を近似した係数のものをガウシアンフィルタと呼ぶ。
- 線形空間フィルタは、後で述べる周波数フィルタとの関係でもわかるように、画像の特定の周波数成分を強めたり弱めることに相当する。
- フィルタ係数の和が1であれば、画像の平均的な明るさは変化しない。1未満であれば暗くなり、1よりも大きければ明るくなる。

# 1. 6 線形空間フィルタ



1/25	1/25	1/25	1/25	1/25
1/25	1/25	1/25	1/25	1/25
1/25	1/25	1/25	1/25	1/25
1/25	1/25	1/25	1/25	1/25
1/25	1/25	1/25	1/25	1/25

5×5の平均化フィルタ

<u>1</u>	<u>4</u>	<u>6</u>	<u>4</u>	<u>1</u>
256	256	256	256	256
<u>4</u>	16	24	16	<u>4</u>
256	256	256	256	256
<u>6</u>	<u>24</u>	<u>36</u>	<u>24</u>	<u>6</u>
256	256	256	256	256
<u>4</u>	<u>16</u>	<u>24</u>	<u>16</u>	<u>4</u>
256	256	256	256	256
<u>1</u>	<u>4</u>	6	<u>4</u>	<u>1</u>
256	256	256	256	256

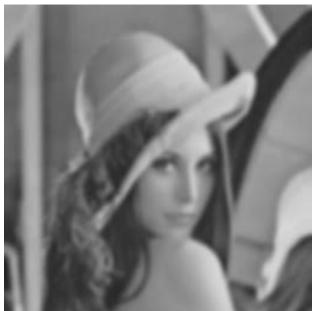
5×5の加重平均化フィルタ (ガウシアンフィルタ)

# 1.6 線形空間フィルタ さまざまなフィルタ

- フィルタ係数にさまざまな値を設定することで、以下の処理を実現可能。
  - 平滑化:画像をなめらかにぼかせる処理
    - ガウス平滑化:係数値がガウス分布にしたがうもの。
    - ボックスフィルタ:係数値が全て等しいもの。平均化フィルタともよぶ。
  - エッジ抽出:画像中の変化を抽出する処理
    - 微分フィルタ:隣接(水平および垂直)画素の差をとる処理。
    - プリューウィットフィルタ、ソーベルフィルタ: 微分フィルタの結果を平滑化することで、ノイズ耐性を強くしたフィルタ。
  - 2次微分:微分を2回繰り返す処理
    - ラプラシアン: 縦方向の2次微分と横方向の2次微分の和。
  - 鮮鋭化、アンシャープマスキング
    - 画像中の高周波成分を増幅することで画像を鮮鋭化できる。このような処理 をアンシャープマスキングともよび、線形空間フィルタで実現可能である。
- これらの処理の詳細については、デジタル画像処理の教科書などを参 照のこと。



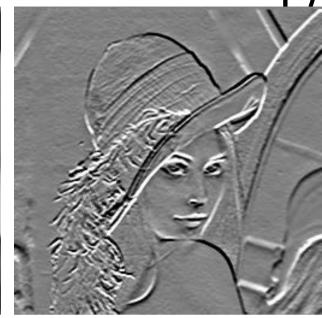
ラプラシアン



ガウス平滑化



鮮鋭化



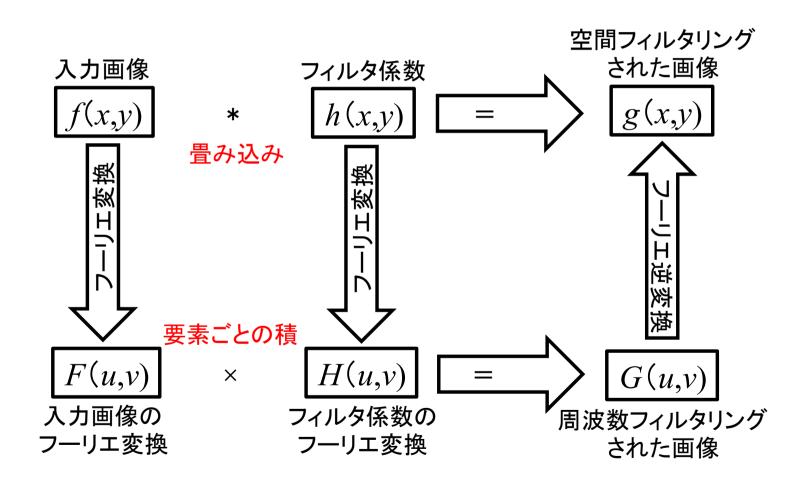
水平プリューウィット

- 1. 元画像は256×256画素
- 2. ガウス平滑化は、11×11のフィルタサイズで標準偏差2のガウス平滑化を施したもの。
- 3. 水平プリューウィット画像とラプラシアン画像は全体の明るさを0.5シフトした。
- 4. 鮮鋭化は、元画像からラプラシアン画像を減算したもの。

### 1.7 空間フィルタと周波数フィルタ

- 線形空間フィルタは入力画像とフィルタ係数の畳み込みである。 フィルタサイズが大きくなるとその計算量は膨大になる。
- 入力画像とフィルタ係数の両方をフーリエ変換し、それらを要素ごとに掛け合わせた結果をフーリエ逆変換すると、線形空間フィルタと同じ結果を得る。このような演算を周波数フィルタリングと呼ぶ。なお、離散フーリエ変換を用いる場合、入力画像とフィルタ係数のサイズは同じにする必要がある。
- 周波数領域でのフィルタリングを用いることで次のメリットがある。
  - フィルタの特性を定量的に設計することができる。
  - 高速フーリエ変換を用いることで、フィルタサイズによらず、計算を 高速に実行することができる。
- 非線形な空間フィルタには、このような単純な関係はない。

# 1.7 空間フィルタと周波数フィルタ









ローパスフィルタ



ハイパスフィルタ



高域強調

- 1. 元画像は256×256画素
- 2. ローパスフィルタは、周波数領域の元画像に対して高周波成分を0にしたのち、空間領域の画像に変換したものである。
- 3. ハイパスフィルタは、周波数領域の元画像に対して低周波成分を0にしたのち、空間領域の画像に変換し、全体の明るさを0.5シフトしたものである。
- 4. 高域強調は、周波数領域の元画像に対して、高周波成分を2倍したのち、空間領域の画像に変換したものである。

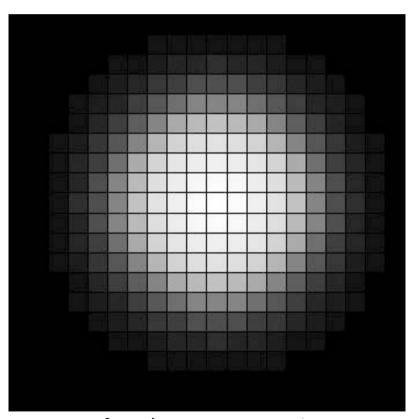
ローパスフィルタは、空間領域での平滑化と類似した結果になる。ハイパスフィルタは、空間領域でのラプラシアンと類似した結果になる。高域強調は、空間領域での鮮鋭化と類似した結果になる。厳密に設計すれば、空間領域での線形フィルタと周波数領域での線形フィルタは同じものになる。

処理後の画像で、特にローパスフィルタ画像とハイパスフィルタ画像に波紋が見える。 これはリンギングなどとよばれる副作用である。画像の周波数特性が急激に変化する 場合に発生しやすい。

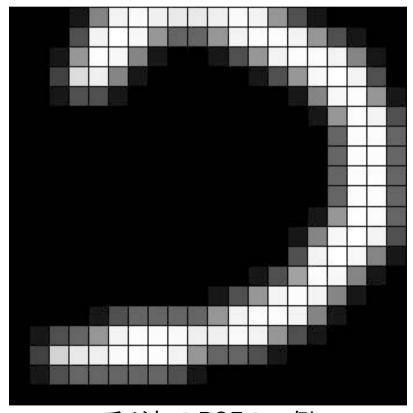
### 1.8 点拡がり関数(PSF)

- 理想的な点光源をカメラで撮影しても、さまざまな要因によって 画像がぼける。このとき、カメラは点像を入力して、ぼけ画像を出 力する装置と考えることができ、その特性を表現したものを点拡 がり関数(PSF: Point Spread Function)と呼ぶ。
- ピンぼけの点拡がり関数をガウス分布によって近似することが 多い。次ページのPSFは関数を画像的に表現したものである。
- ピンぼけに似たぼけ画像を生じる現象に手ぶれがある。手ぶれのPSFは手ぶれ軌跡に対応したものになる。
- PSFは線形空間フィルタの係数によって表現することができる。 従って、ピンぼけや手ぶれによる画像のぼけの発生を線形空間 フィルタによって模擬することができる。

# 1.8 点拡がり関数(PSF)



ピンぼけのPSFの一例

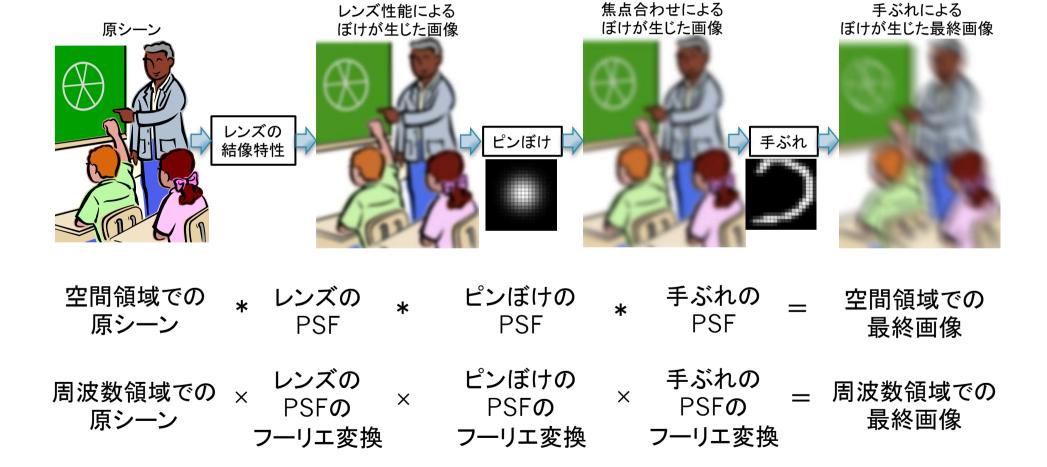


手ぶれのPSFの一例

#### 1.9 線形フィルタの重ね合わせ

- カメラによる撮影では、複数の要因が積み重なって原画像が劣化してゆく。典型的な現象は、レンズ自身の特性によって結像がぼけ、焦点合わせが不完全なためにピンぼけが発生し、さらに手ぶれによって画像がぼける、という状態である。
- 上記のような場合、それぞれのぼけが発生するプロセスを空間線形フィルタによって近似することができる。それらが連続的に作用する現象は、空間領域においては線形フィルタの畳み込みによって表現され、周波数領域においては周波数フィルタの積によって表現される。

## 1.9 線形フィルタの重ね合わせ



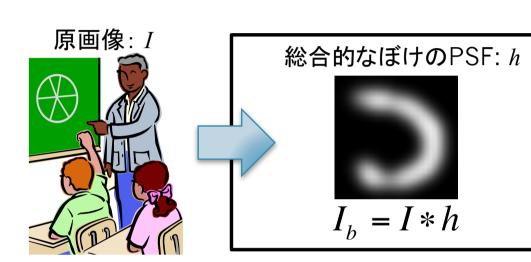
# 1.10 ぼけ画像の復元

- ぼけ画像の生成プロセスを逆にたどることで原画像を復元する処理を 画像復元と呼ぶ。概念的には、PSFの逆関数に相当する演算を、ぼけ 画像に施すことである(この処理をデコンボリューションと呼ぶ)。
- これは一般に不良設定問題になる。従って、適切な拘束条件のもとでの最適化問題に帰着される。数多くの解法が提案されているが、最も良く用いられるのはウィナーフィルタである。
- PSFのフーリエ変換をHとしたとき、周波数領域でのウィナーフィルタは  $H^*$  になる。ここで $H^*$ はHの複素共役、 $\Gamma$ は画像信号のSN(信 $H^2$ + $\Gamma$

号ノイズ比)で決まる定数である。

● 復元の程度はPSFの特性とぼけ画像が含むノイズレベルによって決まる。性質の悪いPSF(例えば、ボックスフィルタと呼ばれる画素近傍を均等に平均化するもの)ではリンギングと呼ばれる副作用が強くでる。また、画像復元は一種の高域強調処理であるため、ノイズを強調する傾向がある。

# 1.10 ぼけ画像の復元



ぼけが生じた最終画像:  $I_h$ 



復元された画像:  $I_{db}$ 



*h-1*: *h*の逆演算に相当する 線形空間フィルタ

$$I_{db} = I_b * h^{-1}$$

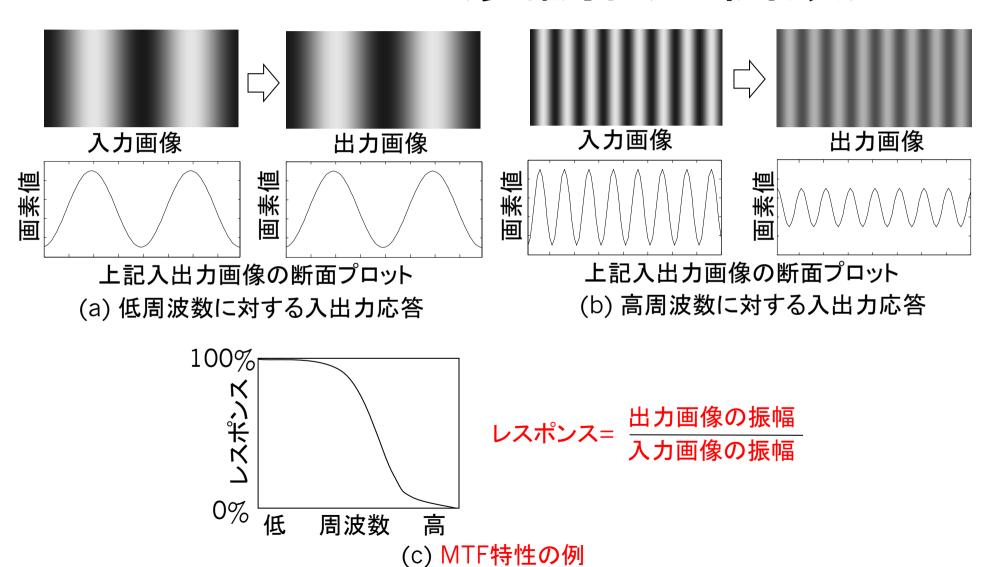
画像の復元

(h-1: デコンボリューションフィルタ)

### 1. 11 MTF(変調伝達関数)

- カメラを2次元波形を入力し2次元波形を出力するシステムと考える。
- MTF(変調伝達関数: Modulation Transform Function) はシステムに正弦波信号を入力したときの、周波数と入出力波形の振幅比率(レスポンス) の特性である。
- カメラのMTFは濃淡値が正弦波状に変化するパターンを観察したとき の入出力関係を、横軸を周波数、縦軸をレスポンスとした特性が用いら れる。
- カメラが撮影するシーンの濃淡変化の周波数が十分に低い場合には、 そのままの状態が再現される。周波数が高くなるにつれレスポンスが 小さくなり、ある周波数以上の濃淡変化は再現することができない。 MTFを用いることで、この性質を定量的に表現することができる。
- MTFは撮影システムの空間周波数領域における振幅特性である。

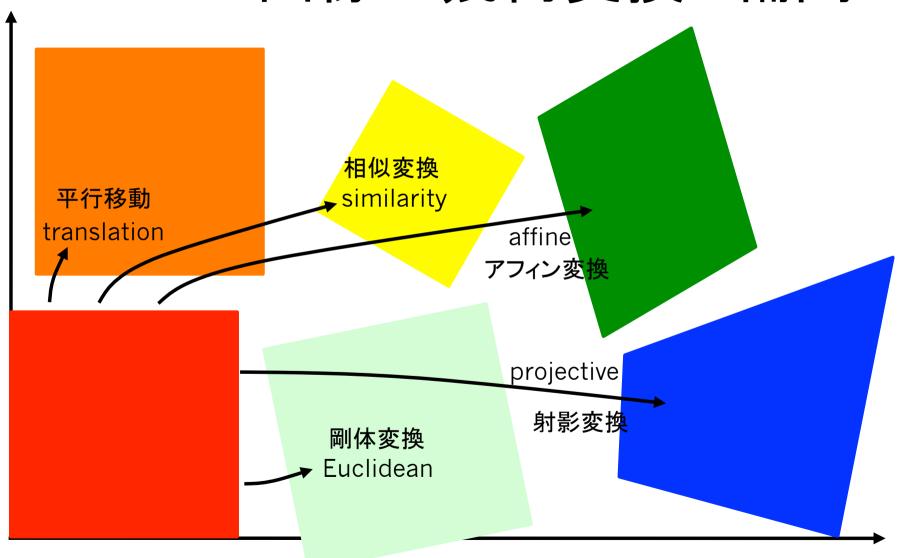
# 1. 11 MTF(変調伝達関数)



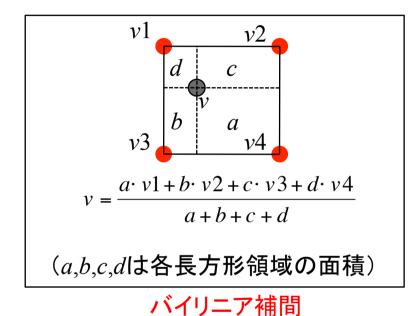
### 1.12 画像の幾何変換と補間

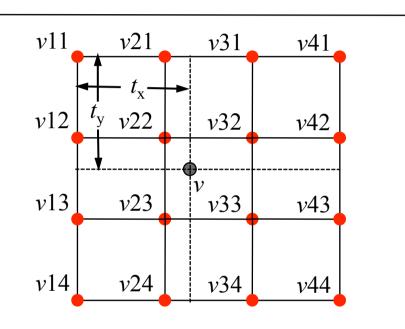
- 画像の幾何変換は画像の形を変形させる処理であり、座標の変換に帰着される。代表的な座標変換は平行移動、 剛体変換、相似変換、アフィン変換、射影変換である。複雑な変形は、画像全体を三角形の領域に分割し、各領域ごとに異なるアフィン変換を適用することなどで実現できる。
- 画像の幾何変換に伴って画素値の補間が必要になる。補間の方法には最近隣補間、バイリニア補間、バイキュービック補間などがある。

# 1.12 画像の幾何変換と補間



#### 1.12 画像の幾何変換と補間





$$h = \begin{cases} |t|^3 - 2|t|^2 + 1 & (|t| \le 1) \\ -|t|^3 + 5|t|^2 - 8|t| + 4 & (1 < |t| \le 2) \\ 0 & (2 < |t|) \end{cases}$$

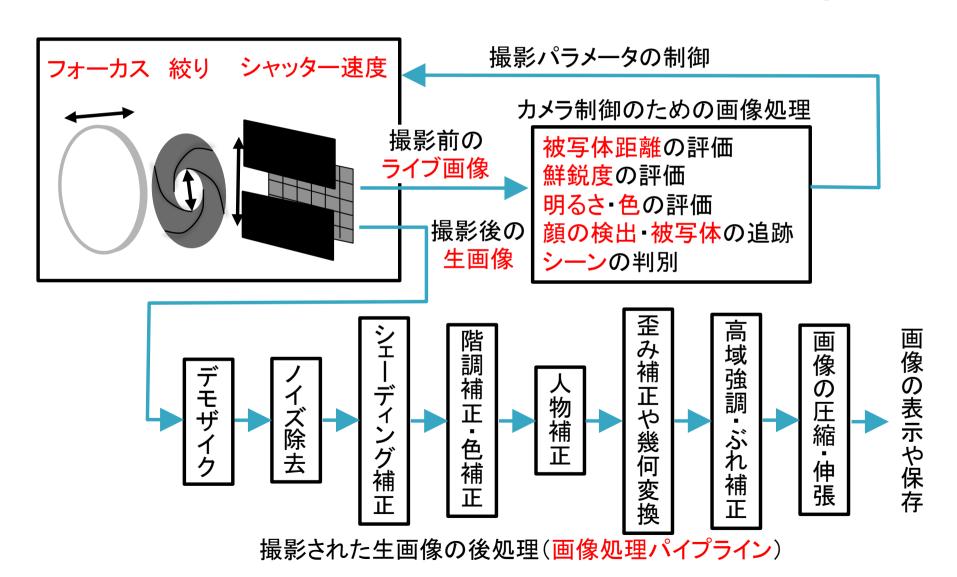
$$h = \begin{cases} |t|^3 - 2|t|^2 + 1 & (|t| \le 1) \\ -|t|^3 + 5|t|^2 - 8|t| + 4 & (1 < |t| \le 2) \\ 0 & (2 < |t|) \end{cases} \quad v = \begin{pmatrix} h_{y1} \\ h_{y2} \\ h_{y3} \\ h_{y4} \end{pmatrix}^T \begin{pmatrix} v11 & v12 & v13 & v14 \\ v21 & v22 & v23 & v23 \\ v31 & v32 & v33 & v34 \\ v41 & v42 & v43 & v44 \end{pmatrix} \begin{pmatrix} h_{y1} \\ h_{y2} \\ h_{y3} \\ h_{y4} \end{pmatrix}$$

# 2章 デジカメの基本機能と画像

# 2. 1 デジカメ画像処理の全体像

- デジカメ画像処理は、大別すると、撮影前のライブ画像に 対して行うものと、撮影後の生画像に対して行うものがある。
- ライブ画像に対して行うものは、ピントや絞りなどの撮影パラメータを決定するために利用する。被写体距離の評価や画像の鮮鋭度の評価、明るさ・色の評価などがある。
- 生画像に対して行うものは、最終画像を生成するための処理である。デモザイクやノイズ除去、階調補正など一連の処理である。画像処理パイプラインと呼ばれる。

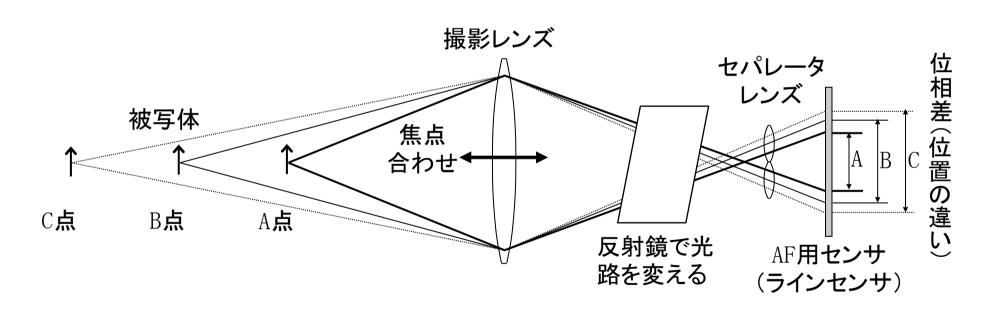
# 2. 1 デジカメ画像処理の全体像



#### 2. 2 位相差AF

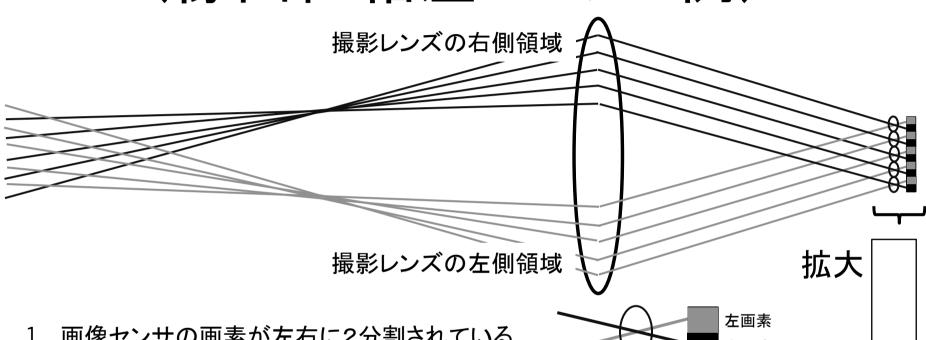
- 撮影状態を制御する項目で基本的なものはフォーカス(ピント)と絞り(露光)である。フォーカスには、手動でピントを合わせるマニュアルフォーカス(MF)と、主要被写体を自動的に判断してピントを合わせるオートフォーカス(AF)がある。
- AFで主に用いられる方式は位相差AFとコントラストAFである。これら以外に、パッシプステレオ方式も利用されている。パッシブステレオ法は2個の画像センサを用いたステレオ計測で、距離を計る方法である。
- 位相差AFは、主に一眼レフで用いられる方式である。撮影レンズでとらえた被写体からの光線の中で、レンズの周辺部を通過する光線をセパレータレンズで抽出し、その像をAFセンサと呼ばれるラインセンサで比較してその位相差(位相差と呼ぶが、実際には像の位置の差である)から被写体までの距離を計算する。
- 通常の位相差AFはライブビューでは動作しない。最近では、画像センサに位相差検出の機能を埋め込んだ<mark>像面位相差AFが実用化され、ライブビューでも位相差AFが可能になっている。</mark>

## 2. 2 位相差AF(通常方式)

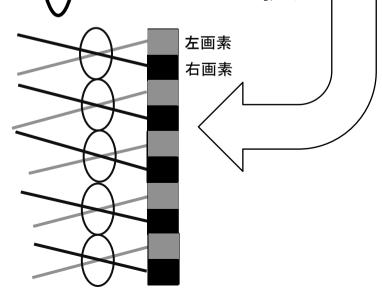


位相差AFの原理

# 2. 2 位相差AF (像面位相差AFの一例)



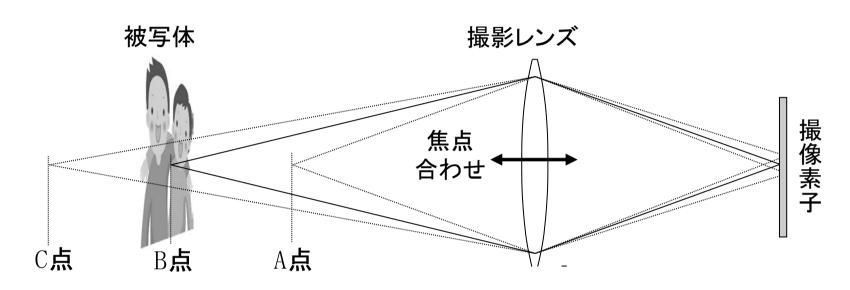
- 1. 画像センサの画素が左右に2分割されている。
- 2. その前に、マイクロレンズが配置されている。
- 3. 灰色の画素が形成する画像は、撮影レンズ の左側領域から観察した像、黒色の画素が 形成する画像は撮影レンズの右側領域から 観察した像になる。
- 4. したがって、この2つの像には被写体までの 距離によって位相差が存在する。



#### 2. 3 コントラストAF

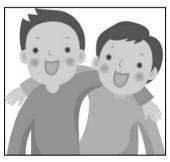
- コントラストAFは、被写体像のコントラスト(像の鮮鋭度)の評価値が最大になる位置に撮影レンズを移動させる方式である。
- この方式は、コントラストが最大になる位置を繰り返し的に 決定する必要があるため、AF動作が遅いとされていた。最 近はAF制御を高速に実行することが可能になったため、改 善されている。
- 一方、被写体像のコントラストを最大にすることができるため、ピントがあった像は鮮明である。

### 2. 3 コントラストAF

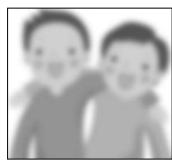




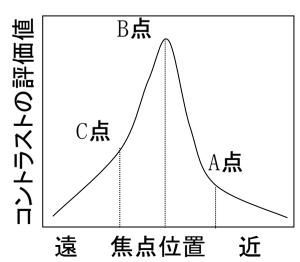
C点に焦点を 合わせた画像



B点に焦点を 合わせた画像



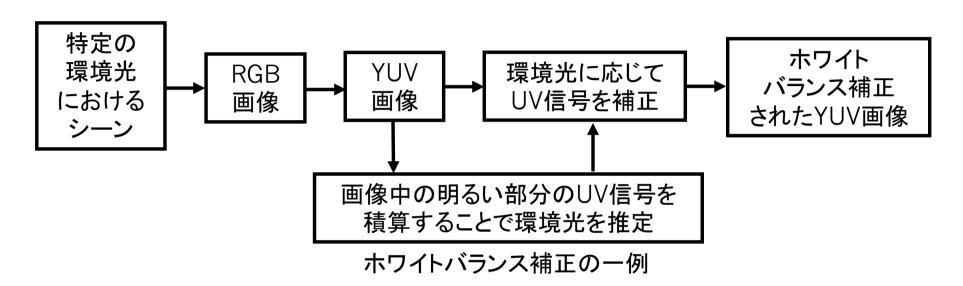
A点に焦点を 合わせた画像



### 2. 4 ホワイトバランス補正

- 人間の視覚は色の恒常性を有するため、電燈・蛍光灯・太陽光など環境光の違いによらず、被写体が持つ本来の色を知覚することができる。これに対してデジタルカメラの画像は環境光の影響を直接受ける。これを補正して、人間にとって自然な画像に色変換する機能をホワイトバランスと呼ぶ。
- ほとんどのデジタルカメラは、ホワイトバランスを自動的に補正するオートホワイトバランス機能を持っている。また、環境光を指定してホワイトバランスを設定することも可能である。
- ホワイトバランス処理の一例は、次ページのように、入力画像をRGB画像から 輝度信号(Y成分)と色差信号(U成分とV成分)で表されるYUV画像に変換する。 次いで、画像中の比較的明るい部分のUV信号を積算した値から、環境光の種 類や色温度(絶対温度Tの黒体から放射される光の色を色温度Tと定義する) を推定する。最後に、推定された環境光に基づいて、UV信号を補正する。この 処理はシーン中の反射率が高い部分に環境光の性質が表れることに基づいて いる。
- YUV信号とRGB信号は3×3の行列によって相互に変換可能である。

### 2. 4 ホワイトバランス補正



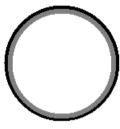
$$\begin{pmatrix} Y \\ U \\ V \end{pmatrix} = \begin{pmatrix} 0.2126 & 0.7152 & 0.0722 \\ -0.1146 & -0.3854 & 0.5000 \\ 0.5000 & -0.4542 & 0.0458 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

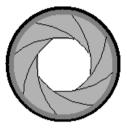
RGB信号とYUV信号の関係

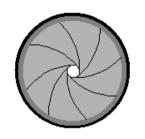
### 2. 5 絞りと画像

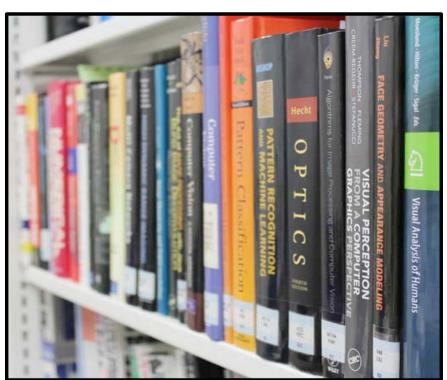
- 絞りは、レンズに入射する光量を調整するための装置で、次ページに示すように、穴の直径が調整できるようになった板状の部品である。
- レンズに入射する光量を自動的に調整する機能を自動露光 (オートアイリス)と呼ぶ。これを実現するために、カメラは画像の明るさを評価している。
- 絞りを開くと像が明るくなるとともに、焦点が合う範囲が狭くなる。 そのような写真は被写界深度が浅いという。
- 絞りを狭めると像が暗くなるとともに、焦点が合う範囲が広くなり、 被写界深度が深くなる。
- 露光時間が一定であれば、絞りを開き、像を明るくするほど、画像のノイズは少なくなる。

# 2. 5 絞りと画像

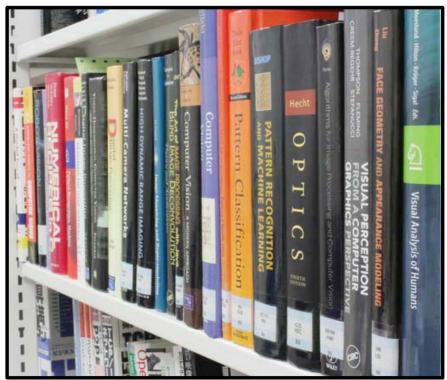








被写界深度が浅い写真

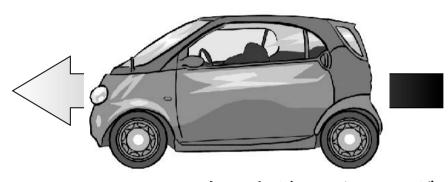


被写界深度が深い写真

### 2. 6 シャッターと画像

- デジタルカメラのシャッターには電子シャッターとメカニカルシャッター(メカシャッター)がある。一般に、レンズ交換式はメカシャッターを用い、レンズー体型は電子シャッターを用いる。
- 電子シャッターにはグローバルシャッターとローリングシャッターがある。
  - CCDはグローバルシャッターで、全画素が同じタイミングで動作する。
  - CMOSはローリングシャッターで、シャッタータイミングが画像の上下でずれる。そのため、動きのあるものを撮影したり、手持ち撮影すると画像が不自然にゆがむ。この歪みをフォーカルプレーン歪みと呼ぶ。

### 2.6 シャッターと画像



全画素が同じタイミングで情景を 撮影するため、像が歪まない



上の画素ほど早いタイミングで 情景を撮影するため、像が歪む

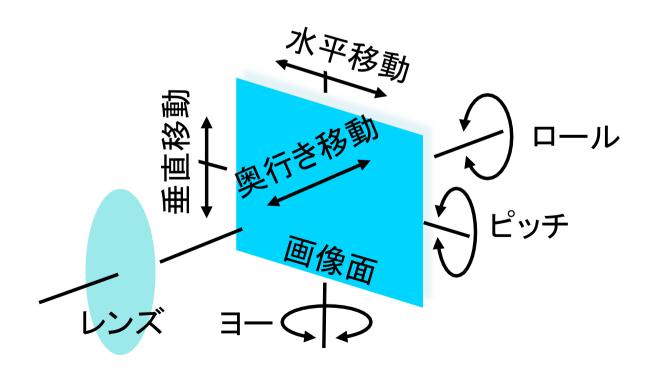


### 2. 7 手ぶれ補正と画像

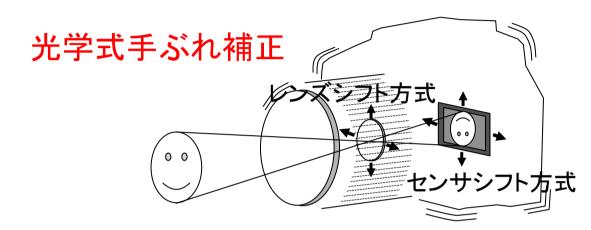
- ぼけた写真の原因は「ピンぼけ」か「手ぶれ」である。手ぶれは撮影中にカメラが動いたことで発生する。手ぶれを補正する方式に光学式と電子式がある。
- 光学式手ぶれ補正は、レンズから撮像部までの光学系を機械的に制御することで、撮像部に対して光像を安定化する方式である。レンズシフト方式とセンサシフト方式がある。
- 電子式手ぶれ補正は、ぶれて撮影されてしまった画像を、 後処理によって修復する処理である。いくつかの方式があ るが、次ページに示す、高速シャッターで連射撮影した複数 の画像を位置合わせして加算合成する方式は代表的なも のである。

### 2. 7 手ぶれ補正と画像

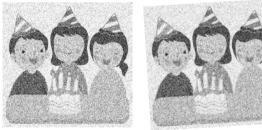
カメラの動きは6つに分解でき、3つの回転(ピッチ、ヨー、ロール)と3つの移動(水平、垂直、奥行き)である。このなかで、ピッチとヨーが支配的。これらに加えて、ロール回転と水平・垂直移動まで補正できる技術が開発されている。



### 2. 7 手ぶれ補正と画像



電子式手ぶれ補正

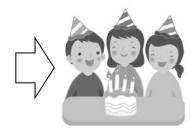






短露光で連射撮影し、ノイズは多いが手ぶれが少ない複数画像

複数画像を位置合わせ・加算合成

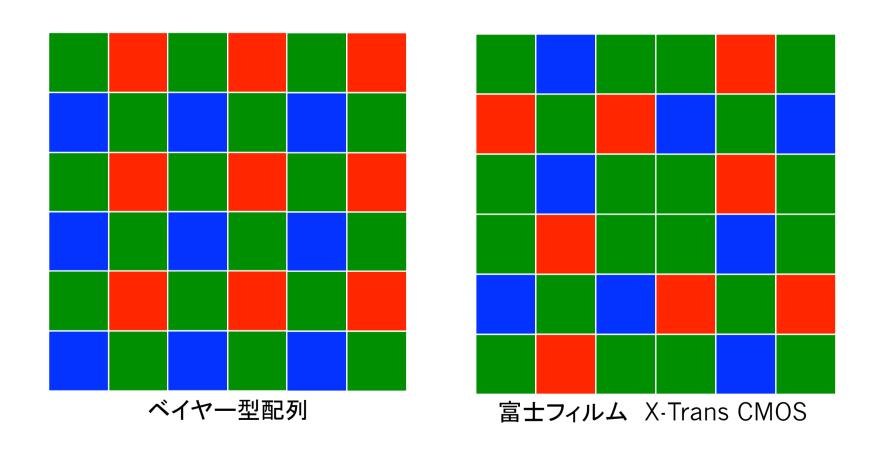


# 3章 デジカメ画像処理の各論

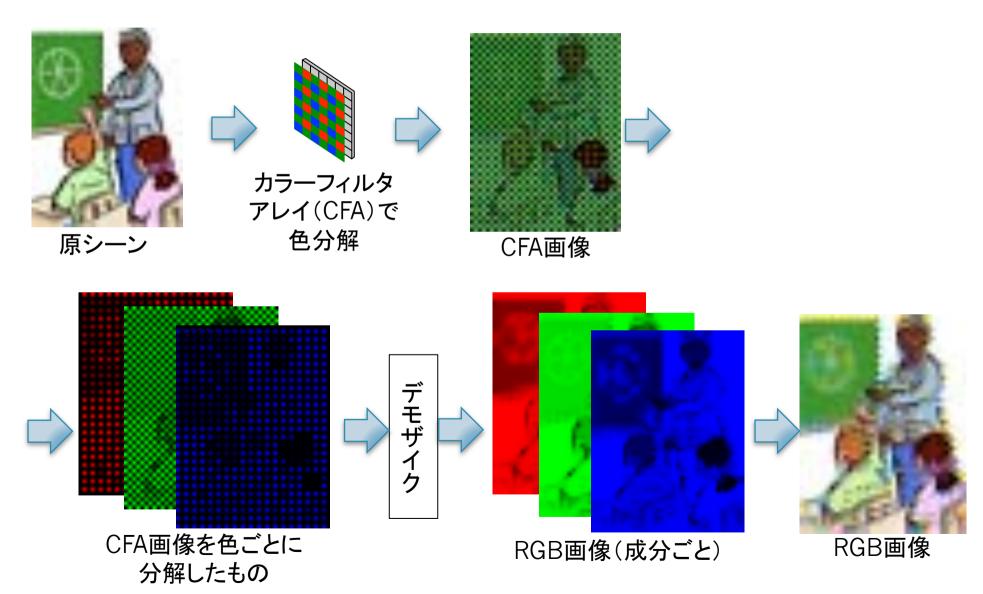
### 3. 1 カラーフィルタアレイとデモザイク

- カラー画像の撮影は、単板式でカラーフィルタアレイを用いて色分解する方法が一般的である。その他に、3板式と、単板で直接3原色を取得する方式注がある。
- カラーフィルタアレイで色分解した画像(CFA画像)は、各画素がRGBいずれか一つの値だけを持つ不完全なカラー画像である。
- CFA画像からRGB画像を生成する処理をデモザイクと呼ぶ。
- デモザイクは画像信号の補間である。しかし、バイリニア法やバイキュービック法など、画像拡大に利用する方法を用いると、エッジなど信号が急激に変化する部分で「偽色」と呼ばれるアーチファクトが発生する。
- デモザイクに適した画像の補間法が開発されている。RGB信号の相関を利用した補間法などにより、解像度をそれほど犠牲にすることなく、偽色を抑えることが可能である。

### 3. 1 カラーフィルタアレイとデモザイク



### 3. 1 カラーフィルタアレイとデモザイク



### 3. 2 画像に発生するノイズ

- 画像に発生するノイズは画素値がランダムにあるいは固定的に、望ましい数値から変化することである。その原因は、フォトダイオードの暗電流による固定パターンノイズ、フォトダイオードに近接したアンプのノイズ、画素をリセットした時に発生するノイズ、電荷を転送するときに発生する回路的なノイズなどである。
- これらのノイズは、半導体プロセスの改良や回路的な工夫により改善されているが、ゼロにすることはできない。従って、画像処理によってノイズを除去することは重要である。

## 3.2 画像に発生するノイズ









ノイズが無い画像

ランダムノイズ

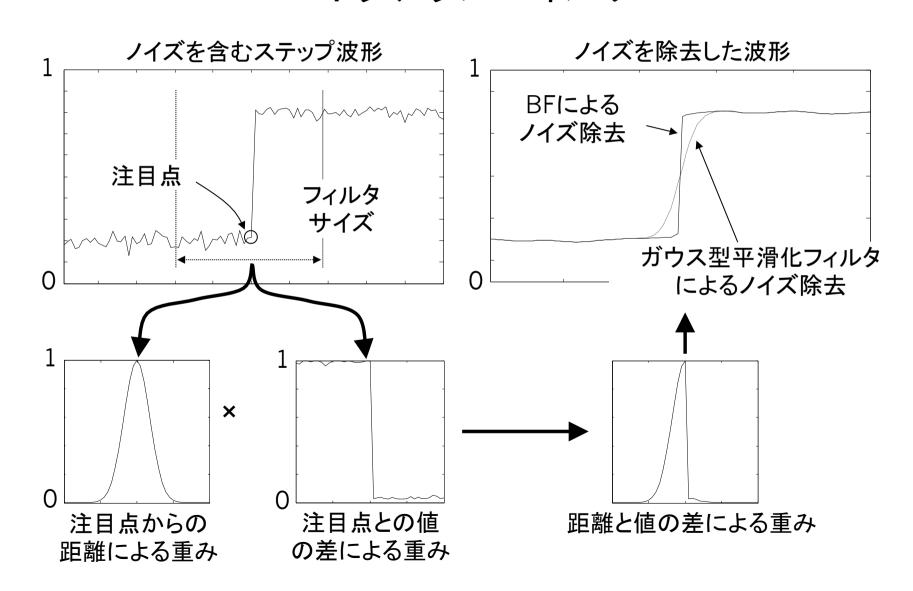
固定パターン的 インパルスノイズ

両ノイズが 重畳した画像

# 3.3 ノイズ除去ーバイラテラルフィルター

- ランダムノイズに対する古典的な除去手法は、注目画素の周辺に矩形や円形の小領域を設定し、その領域の画素値を加重平均するものである。インパルスノイズに対しては、領域の中央値を出力するメディアンフィルタがよく利用される。
- ノイズ除去は、原シーンに存在する情報を保存しつつ、ノイズを除去する必要がある。特に、エッジの保存が重要である。ここでは、代表的なエッジ保存型平滑化フィルタであるバイラテラルフィルタ(BF)を紹介する。
- 簡単のために1次元信号で説明する。ノイズを含むステップ信号の直前の点を注目点とする。注目点からの距離に応じたガウス分布重みと、注目点との値の差に応じたガウス分布重みの、2つの重み係数を設定する。これを掛け合わせることで、注目点からの距離と、値の差の両方を考慮した重み係数を決定する。これを用いて、近傍画素を重み付き平滑化することでノイズを除去する。BFによるノイズ除去はガウス型平滑化によるノイズ除去に比べて、エッジを保存する性能が優れている。

# 3.3 ノイズ除去ーバイラテラルフィルター



### 3. 4 PSNRによる画質評価

- ノイズ除去や画像圧縮の性能を比較する時、2画像の違いを客観的に評価する必要がある。この評価指標にPSNR(Peak Signal-to-Noise Ratio)、RMS(Root Mean Square)、MSE(Mean Square Error)などがある。
- PSNRは最もよく用いられる数値で原画像をS(i,j)、処理後の画像をN(i,j)とするとき、

$$PSNR = 10 \log \left( 255^{2} / \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} \left( S(i,j) - N(i,j) \right)^{2} \right)$$

によって計算される(m,nは縦横の画素数、単位はdB)。

● 2画像が完全に同じであればPSNRは無限大になり、40dB以上でほぼ区別がつかないとされる。

### 3. 4 PSNRによる画質評価

入力画像



白色ガウス雑音 を加える PSNRで比較する

ノイズを加えた画像



処理後の画像



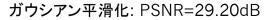
雑音を除去

### 3. 4 PSNRによる画質評価











バイラテラルフィルタ: PSNR=29.92dB



BM3D[1]: PSNR=32.05dB

### 3.5 シェーディング補正

- デジカメ画像におけるシェーディングは、画像の周辺部分が暗くなる現象である。似た現象にヴィネットがある。
- レンズの光軸に対して光線が角度 $\theta$ で入射する場合、像の明るさが $\cos \theta$  の4乗に比例して暗くなる。これがシェーディングの主な原因である。同じ現象が、撮像部のマイクロレンズでも発生する。
- ヴィネッティングはレンズ鏡筒のふちで光線の一部が遮られることによって周辺が暗くなる現象である。
- シェーディングはレンズと撮像素子の組合せに対して固定的であるから、画素位置に応じた係数で画素値を補正することが可能である。

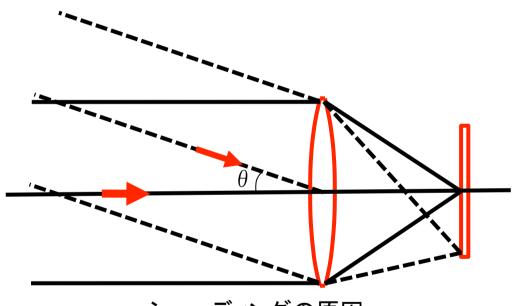
# 3.5 シェーディング補正



周辺光量が低下した画像



周辺光量の低下を補正した画像

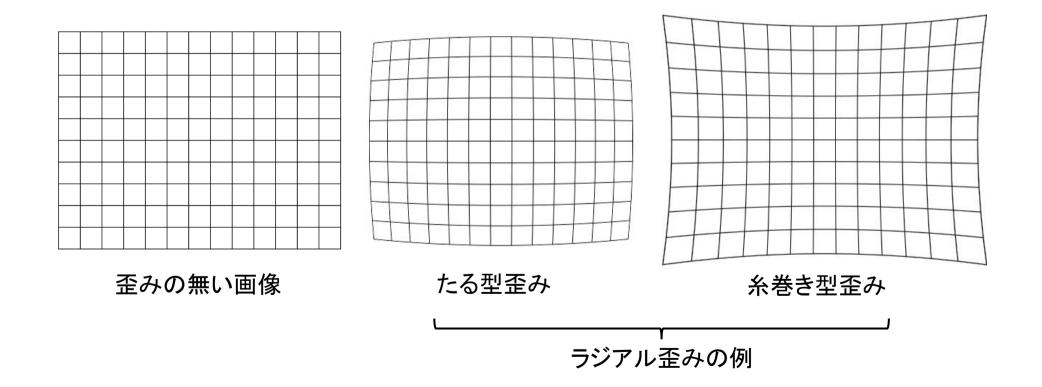


シェーディングの原因

### 3.6 歪みの補正

- 撮影された画像が、幾何学的に変形する現象を歪みと呼ぶ。歪 みの発生には主に2つの原因がある。
- 第一の原因は、レンズの屈折が中央から離れるに従って、理想 状態から外れることである。これによる歪みをラジアル歪み(半 径方向歪み)と呼ぶ。ラジアル歪みが発生した画像の例を次 ページに示す。
- 第二の原因は、レンズと撮像素子が平行に取り付けられていないことによるものである。これによる歪みをタンジェンシャル歪み(円周方向歪み)と呼ぶ。
- 画像に生じた歪みは、画素位置に応じた変位量を与えて補正することができる。

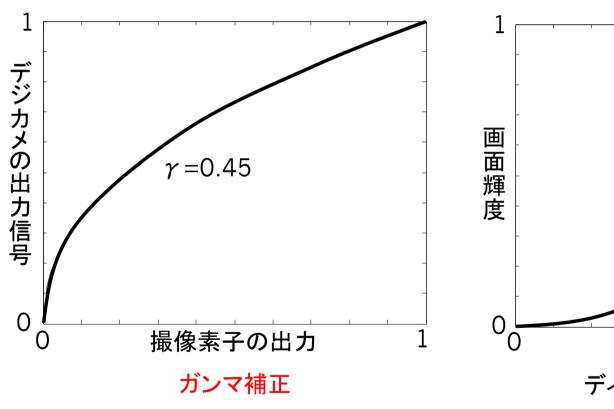
## 3.6 歪みの補正

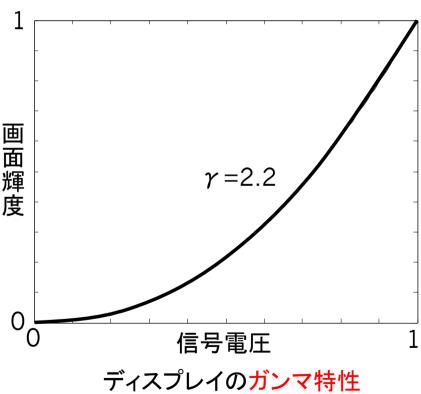


### 3. 7 ガンマ補正

- パソコンやテレビのディスプレイは、入力信号の2.2乗に 比例した明るさで表示されるガンマ特性を持っている。これ は、標準的なディスプレイとして利用されていたブラウン管 の表示特性である。LCDなど新しいディスプレイ技術は、そ の技術固有の発色特性を有するが、過去のシステムとの 互換性を保つためにガンマ特性を持たせている。
- 撮像素子は像の明るさに比例した信号を出力する。デジタルカメラ内部でガンマ補正と呼ばれる階調補正を加えることで、撮影されるシーンをディスプレイ上に忠実に再現する。

## 3. 7 ガンマ補正

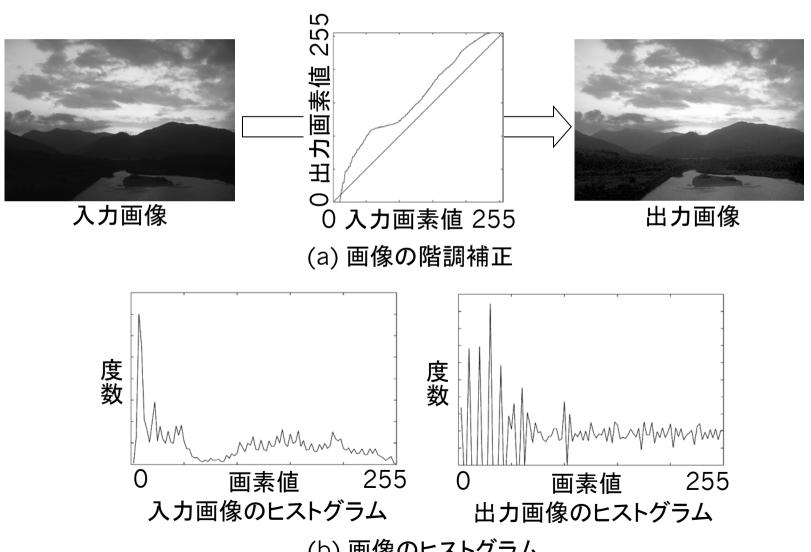




### 3.8 トーンマネジメント

- ガンマ補正を含めた、カメラ内で行われるさまざまな階調特性の制御をトーンマネジメントと呼ぶ。トーンマネジメントは、シーンの明暗範囲とカメラで撮影可能なダイナミックレンジが大きく異なる場合や、人間にとって好ましい画像再現を実現する場合などに重要である。
- 代表的なトーンマネジメントにヒストグラム平坦化と呼ばれる処理がある。ヒストグラムとは、画像における画素値とその度数の関係である。ヒストグラム平坦化処理では、画像全体のヒストグラムを平坦化することで、画素値が密集した明暗範囲を拡大し、その視認性を改善する。ただし、必ずしも好ましい画像再現になるとは限らない。

## 3.8 トーンマネジメント



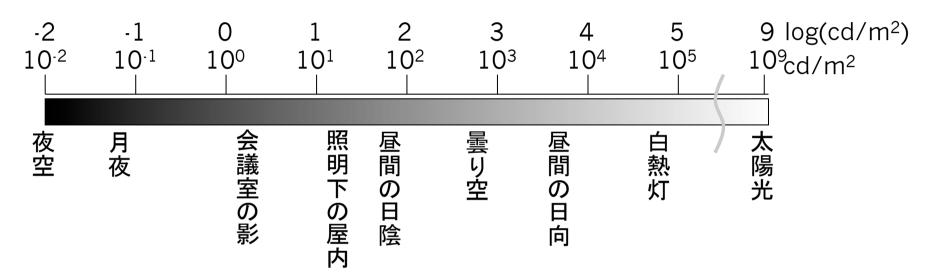
(b) 画像のヒストグラム

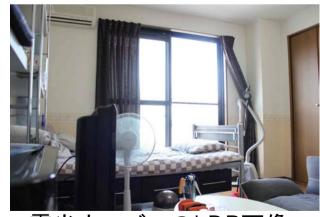
# 3. 9 高ダイナミックレンジ (HDR)画像

- デジタルカメラの撮像素子のダイナミックレンジ(同時に撮影可能な明暗の範囲)は50~70dB(明暗の対数スケールで2.5~3.5桁)である。これに対して、実シーンの明暗の範囲は100dB(5桁)以上になることもある。
- 通常の撮影では不可能な、ダイナミックレンジが広い画像を高ダイナミックレンジ(HDR)画像と呼ぶ。
- 通常のデジタルカメラでHDR画像を取得する場合、同じシーンに対して露光量を変化させた複数の画像を撮影し、それらを合成することでHDR画像を生成する。
- HDR画像を生成しても、通常のディスプレイではそのまま表示できない。その場合、トーンマッピングと呼ばれる一種の階調変換によってダイナミックレンジを圧縮した後、画像を表示する。

#### 69

# 3. 9 高ダイナミックレンジ (HDR)画像





露光オーバーのLDR画像

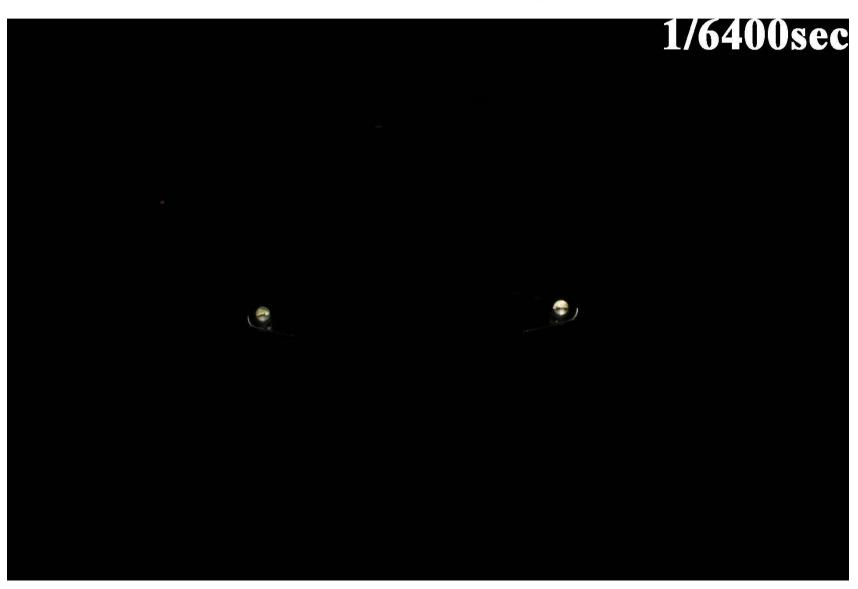


露光アンダーのLDR画像

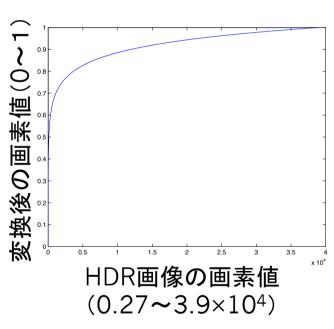


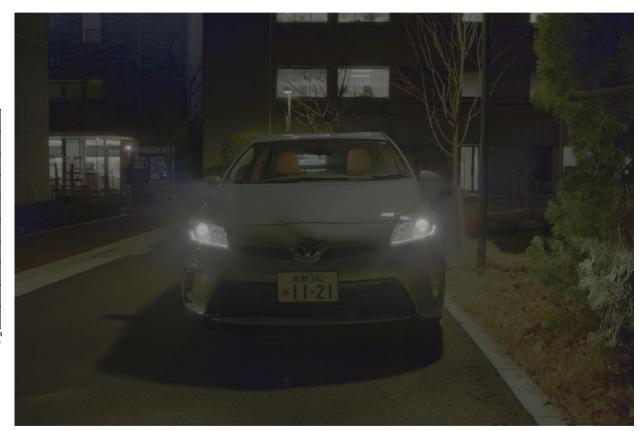
多重露光画像を合成した画像

## HDR画像の元画像の一例



## トーンマッピングで変換した画像



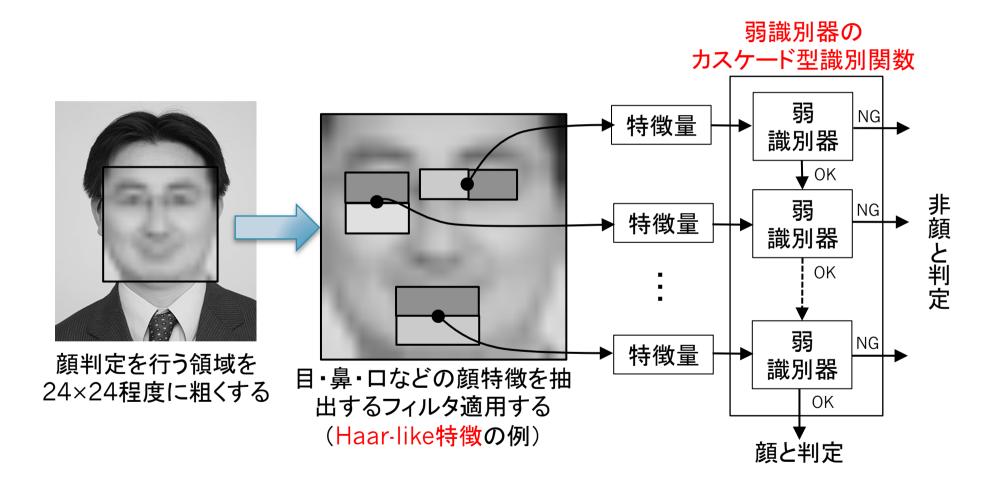


HDR画像をトーンマッピングで256階調に変換した画像

#### 3.10 顔画像の検出

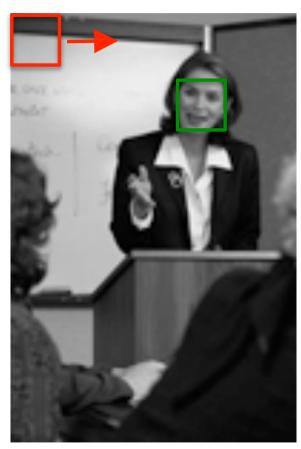
- 顔検出機能はほとんどのデジタルカメラに実装され、顔へのオートフォーカスやオートアイリス(自動露光)に利用されている。
- 顔を検出するアルゴリズムの概要は次のようなものである。
  - 顔判定を行う領域は24×24画素程度の大きさ固定の小領域である。
  - 小領域に対して、目・鼻・口などの顔部品を抽出する空間フィルタを 適用する。空間フィルタの設計に各社のノウハウがある。初期におい てはHaar-like特徴が用いられていた(次ページで図示したもの)。
  - 小領域から多くの特徴量が抽出され、それらを識別関数に入力することで顔の判定を行う。識別関数は弱識別器と呼ばれる線形識別関数をカスケード接続するものが有名である。
  - 入力画像からもれなく顔を検出するために、入力画像を段階的に縮小したピラミッド画像を作成する。そして、各階層の画像に対して小領域をラスター走査することで網羅的に探索する。

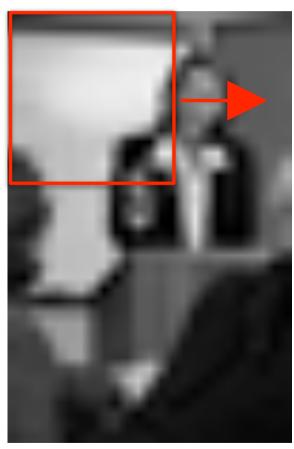
#### 3.10 顔画像の検出



#### 3.10 顔画像の検出







入力画像を段階的に縮小した複数の画像に対して、固定された大きさの小領域(24×24画素程度)の顔判定をラスター走査的に網羅的に実施することで、画像中の顔を検出する

#### 3. 11 さまざまな顔画像処理

- 顔検出だけではなく、次のような処理が可能である。
  - 年齢・性別の推定や、個人の同定
  - 顔の表情(特に笑顔度)の推定
  - 目つむりの判定
  - ◆ 赤目の補正、瞳補正(目を大きくする、瞳に光を入れる)
  - 美肌処理(しみ・そばかす・肌荒れ・毛穴の軽減)や 美白処理
  - 小顔補正(顔を小さくしたり、細くする)

#### 3.12 移動物体の追跡

- 動被写体の追跡は、乗り物・スポーツ・ペットなどを撮影するシーンで、前後左右に動く被写体に連続的にフォーカスを合わせ続ける場合に必要である。これを画像処理で実現するには、指定した動被写体領域の画像特徴を、連続する画像フレームで探索する。
- 安定した追跡を実現するには、形状変化や背景領域の変化に頑健な画像特徴を用いる必要がある。領域の画素値をそのまま用いる方式(テンプレートマッチング)は 形状変換に弱い。
- しばしば用いられるのは、カラーヒストグラムである。カラーヒストグラムは画素のカラー値を類似した色ごとにグループ化し、グループごとの度数を計数した特徴量である。カラーヒストグラムを全度数で正規化した正規化カラーヒストグラムは物体の形状変化に頑健な特徴である。
- 連続するフレーム間での画像特徴のマッチングを確率的な事象と考えることで、より 頑健な追跡を実現することができる。このような確率的手法にベイジアンフィルタが ある。この手法では追跡状態(位置)とその確率を関係づけるので、最適な追跡結 果を求めることや、複数の仮説を同時に処理することができる。具体的なアルゴリズ ムにカルマンフィルタやパーティクルフィルタがある。

# 3. 12 移動物体の追跡 カラーヒストグラム

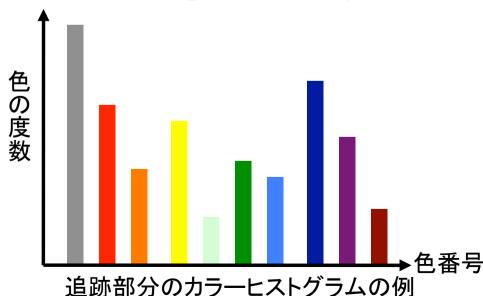


追跡部分のテンプレート画像

動画の中で移動する物体を、画像処理で追跡することは難しい問題である。なぜなら、

- 1. 移動物体の形・大きさが変化する
- 2. 背景が変化する

単純なカラー画像のテンプレートマッチングでは動作が安定しない。形状変化に強いカラーヒストグラムを用いる手法が有効である。



#### 3. 12 移動物体の追跡 カラーヒストグラムの類似度

- ◆ カラーヒストグラムを用いる場合、カラーヒストグラム間の類似度を計算する必要がある。
- 類似度の算出には、ヒストグラムインタセクションとバタチャリヤ係数がよく利用される。p<sub>i</sub>、q<sub>i</sub>をそれぞれの正規化カラーヒストグラム(ただし、i=1~Nは色番号)としたとき、
  - ヒストグラムインタセクション:  $\rho(p,q) = \sum_{i=1}^{N} \min(p_i,q_i)$
  - バタチャリヤ係数:  $\rho(p,q) = \sum_{i=1}^{N} \sqrt{p_i q_i}$
- これらの式は、ヒストグラムの分布が類似しているほど値 が大きくなり、最大値は1である。

## 3. 12 移動物体の追跡 ベイジアンフィルタ

- 動画像の画像列を $I_1,...,I_t$ 、そこにある移動物体の位置を $x_1$ , ..., $x_t$ とする。 $p(x_t|I_{1:t})$ は、画像列 $I_1,...,I_t$ を観察した上で、移動体の位置が $x_t$ である条件付確率とする。
- ベイジアンフィルタは、1時刻前の条件付確率 $p(x_{t-1}|I_{1:t-1})$ に、以下のように予測と観測・修正を繰り返す方法である。

  - 観測•修正:  $\rho(x_t|I_{1:t}) = \rho(I_t|x_t)\rho(x_t|I_{1:t-1})$
- カルマンフィルタは、予測・観測のモデルが線形で、ノイズ がガウス分布にしたがう場合のベイジアンフィルタである。
- パーティクルフィルタは、それ以外の場合をモンテカルロ法 で近似計算する手法である。

#### 3. 13 超解像処理

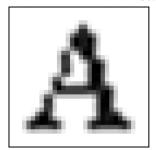
- 通常の画像の拡大は標本化定理に基づいて実行される。具体的な方法に最近隣補間、バイリニア補間、バイキュービック補間などがあるが、いずれも元画像に存在しない周波数成分を発生させる能力はない。
- 超解像処理は、単純な拡大と異なり、元画像に存在しない高周波成分を生成することで、より詳細な画像表現を目指すものである。複数画像からの超解像処理と、1枚の画像からの超解像処理がある。
- 複数画像からの超解像処理は、サブピクセル精度で位置合わせした画像群に画像 復元処理を行うことで画素数を増やす処理である。複数の画像を、元シーンの高周 波成分の折り返しが残るように撮影し、それらを位置合わせした後に、高周波成分 を強調する。
- 1枚の画像からの超解像処理は、標本化定理の枠組みを外した処理で、その多くは「学習ベース超解像」と呼ばれるものである。この手法は多量の学習画像を用いて、低解像度画像と高解像度画像の対応をモデル化(学習)する。そして、未知の低解像度画像に対して、このモデルを参照して高画素化する。
- 合志が提案する「<mark>非線形信号処理</mark>」による超解像は、画像中の高周波成分を抽出し、 その3乗を原信号に加える方式である(1枚の画像からの超解像に分類される)。従 来の超解像に比べて処理が軽く、汎用性もあり優れている。

#### 3. 13 超解像処理

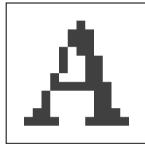
バイキュービック補間







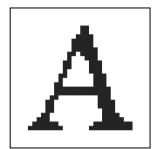
最近隣補間



(a)補間による拡大 高画素化されるが、元画像に存在しない高周波 成分は発生しないため、画像がぼける。







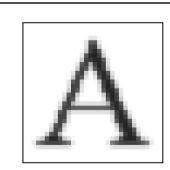
(c) 1枚からの超解像 画像は鮮鋭化されるが、予想でき ない結果になる場合もある。







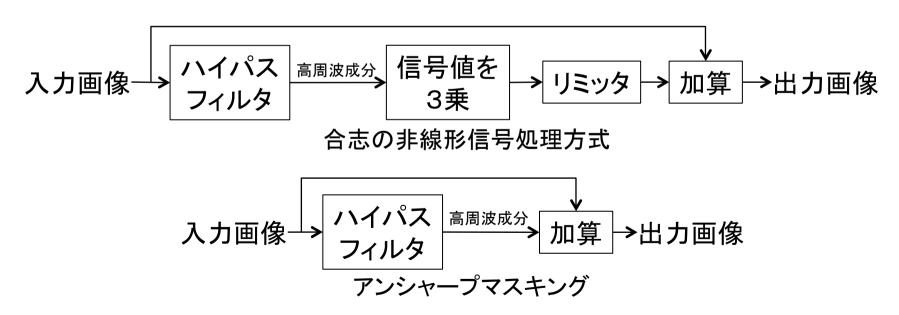




(b) 複数画像からの超解像 折り返し歪みを含む複数の画像を、正確に位置合わせ したのち、その中の高周波成分を増幅する処理。

#### 3.13 超解像処理 合志の非線形信号処理方式1/2

● アンシャープマスキングは、入力画像中の高周波成分を線形に強調する。合志の方法[1]は非線形(3乗)に強調する。その結果、強い高周波成分がより強められ、弱いものはそれほど強められない。また、三角関数の3倍角定理から、抽出した高周波の3倍の周波数が発生する。



<sup>[1]</sup> S. Goshi, "A New Signal Processing Method for Video –Reproduces the Frequency Spectrum Exceeding the Nyquist Frequency-," MMSys'12 Proceedings of the 3<sup>rd</sup> Multimedia Systems Conference, pp.47-52, 2012.

#### 3.13 超解像処理 合志の非線形信号処理方式1/2

合志の方法



濃淡変化が小さい部分で自然な画像が 表現できている

アンシャープマスキング

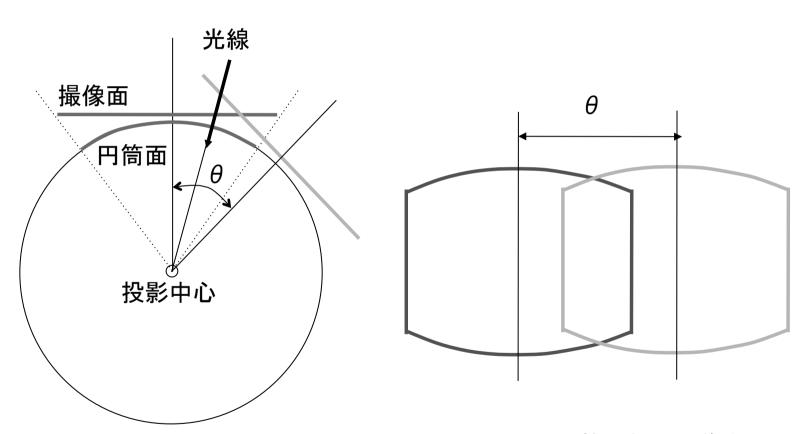


ノイズが強調され、画像が荒れる

#### 3. 14 パノラマ画像

- 1枚のデジカメ画像で撮影できる範囲はレンズの視野角によって制限される。 魚眼レンズなどの特殊なレンズを用いれば180度の視野を得る ことも可能である。また、パノラマ画像合成技術を用いることで、標準的 なレンズを装着した状態でも、広視野画像を獲得することが可能になる。
- パノラマ画像合成技術は、重なりを持つように撮影した複数の画像をつなぎ合わせ、1枚の広視野画像を生成する。
- 複数の画像を撮影するときに、三脚などを用いることでカメラの投影中心が移動しなければ、奥行きの深いシーンなどいかなるシーンでも、適当な投影面を設定することで、継ぎ目なく合成ことが可能である。
- 手持ち撮影などカメラが移動する場合に継ぎ目なく合成できるのは、厳密には平面的な被写体に限られる。
- パノラマ画像合成では、撮影された画像を仮想投影面に射影し、変形を受けた画像どうしを合成する。仮想投影面には円筒面、球面、平面などがあり、投影面の選択によって合成画像の仕上がりが異なる。

## 3. 14 円筒面投影による パノラマ画像合成



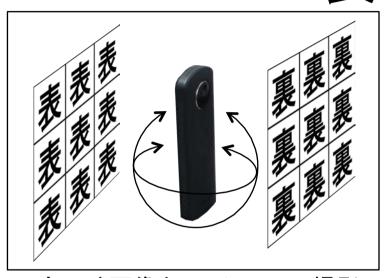
Step1: 撮像面を円筒面へ投影する

Step2: 円筒面投影画像を 位置合わせして合成する

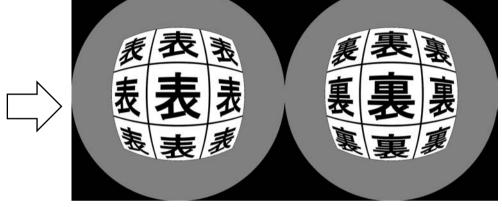
#### 3. 14 全天球画像

- リコーTHETAの登場によって、全天球を表現するパノラマ画像を 撮影することが容易になった。
- 1枚の無限レンズで撮影する画像は円形である。通常、円の中心から像までの距離が、光線の入射角に比例する等距離射影方式で像が形成される。
- THETAでは、表裏に対応した2枚の円形像を、一枚の全天球画像に変換する。その全天球画像は、正距円筒図法で描かれた円筒面画像である。
- 円筒面画像は広い視野を確認することに適しているが、シーン中での直線が湾曲するなどの歪みがある。これを、通常の透視投影画像に変換することが可能である。

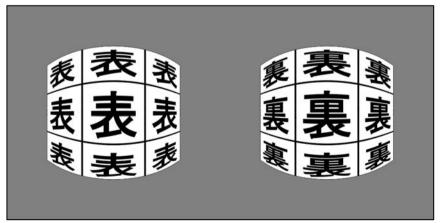
# 3.14 全天球画像の撮影から 87 表示まで



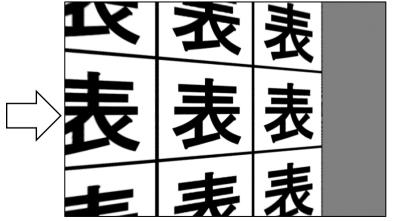
全天球画像をワンショットで撮影



魚眼レンズによる生画像



正距円筒画像



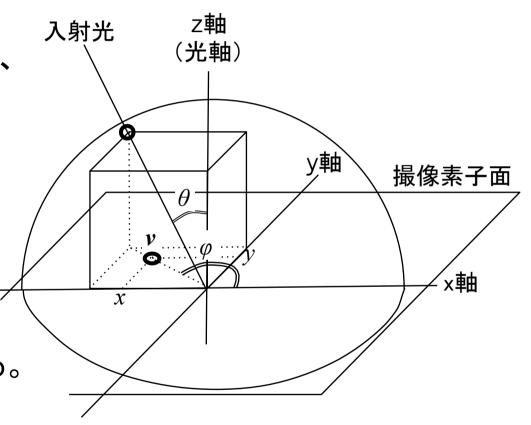
視野の一部を透視投影した画像

#### 3. 14 魚眼レンズによる撮影

 等距離射影モデルでは、 入射光が光軸となす角θ、 光軸まわりの回転角φの とき、像の位置vは

$$\mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix} = f\theta \begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix}$$

となる。ここで f は撮影 範囲を決める定数である。



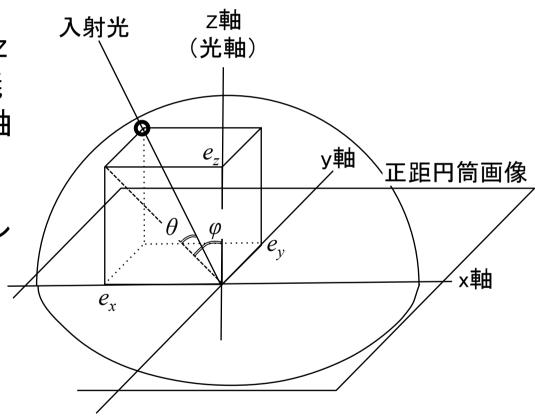
#### 3. 14 正距円筒図法

 光線方向をy軸まわりの 回転角(方位角)φと、x-z 平面からの仰角θで定義 し、画像のx軸をφに、y軸 をθに対応させる。

● 光線方向の単位ベクトル

は

$$\begin{pmatrix} e_x \\ e_y \\ e_z \end{pmatrix} = \begin{pmatrix} \sin\varphi\cos\theta \\ \sin\theta \\ \cos\varphi\cos\theta \end{pmatrix}$$



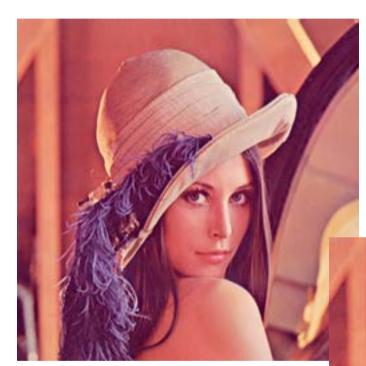
となる。

### 4章 画像映像信号の圧縮技術

#### 4. 1 JPEGによる静止画の圧縮

- 画像や映像の生データの容量は非常に大きい。例えば、1,790 万画素の画像を非圧縮で保存するRAW形式のファイル容量の 一例は24.5MBである。
- デジタルカメラは静止画をJPEGで符号化することで、ファイル容量を5分の1程度に圧縮している(RAW形式のファイル容量に対して)。
- 次ページは256×256画素のRGB画像と、それをJPEG圧縮した例である。この場合、原画像のファイル容量196KB(=256×256×3)に対して、10分の1程度に圧縮しても顕著な劣化が見られない。20分の1の圧縮では画質劣化が目立つ。
- 圧縮率が高いJPEG画像に見られる画質劣化は、ブロック境界がはっきりと見える「ブロック歪み」と、画像のエッジ近傍に現れる「モスキートノイズ」と呼ばれる細かい揺らぎである。

#### 4. 1 JPEGによる静止画の圧縮



原画(196KB)



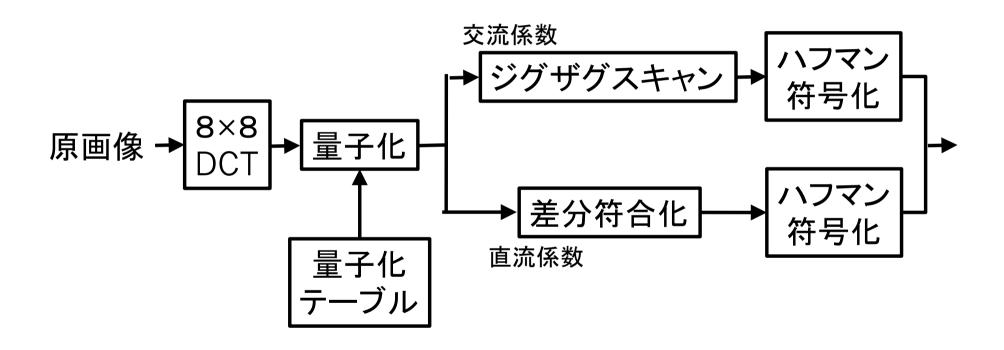


JPEG圧縮(8KB)

#### 4.2 JPEG符号化器の構成

- JPEG符号化器の構成を次ページに示す。
- 画像を8×8の小画像に分解し、これを2次元離散コサイン変換(DCT) する。DCTは一種の周波数領域への変換であり、空間領域の画像(64個の画素値)を1個の直流成分と、63個の交流成分に変換する。
- DCTで計算される64個の成分の中で直流成分に最も多くのビットを割り当てる。交流成分については、低周波成分ほど多くのビットを割り当て、高周波成分ほど少ないビットを割り当てる。
- 63個の交流成分は、「ジグザクスキャン」と呼ばれる方法で、水平・垂 直周波数の順序に従って1次元化する。
- 直流成分は隣り合うブロック間で相関が高いため、差分を符号化する。
- 最後に「ハフマン符号化」を適用する。

#### 4.2 JPEG符号化器の構成



DCT: 離散コサイン変換(Discrete Cosine Transform)

ハフマン符合:発生確率の高いシンボルに短い符号を割り当てる基本的な符号化法

#### 4.3 2次元DCTの計算手順

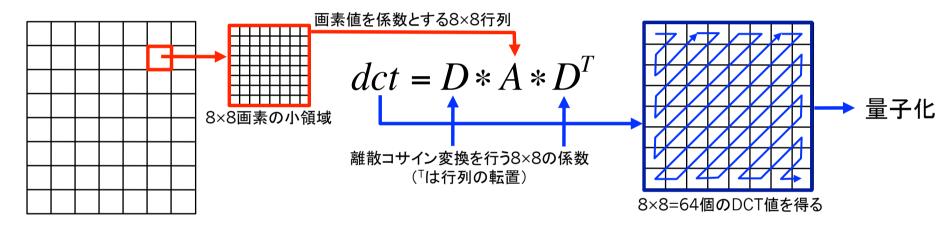
- JPEG符号化の最重要部分が2次元DCTである。JPEGでは8×8画素単位で2次元DCTを行う。
- 2次元DCTの計算手順は、画素値を係数とする行列(次ページのA)にDCT変換行列(次ページのDとD<sup>T</sup>)を乗算する演算になる。変換行列の係数は、第一行が画像の直流成分への変換を意味し、第2行から第8行にしたがって高周波成分への変換を意味する。
- 2次元DCTで8×8のDCT値が計算される。DCT値の第1列・第1行の要素が直流成分、それ以外が低周波から高周波までの交流成分である。DCT値は図のようにジグザグスキャンすることで1次元化し、次いで、量子化する。

#### 4.3 2次元DCTの計算手順

入力画像を8×8画素 の小領域に分割

小領域を単位として 2次元離散コサイン変換(DCT)を施す ジグザグスキャンし量子化

DCT値を



$$D = \begin{pmatrix} 0.353 & 0.353 & 0.353 & 0.353 & 0.353 & 0.353 & 0.353 & 0.353 \\ 0.490 & 0.415 & 0.277 & 0.097 & -0.097 & -0.277 & -0.415 & -0.490 \\ 0.461 & 0.191 & -0.191 & -0.461 & -0.461 & -0.191 & 0.191 & 0.461 \\ 0.415 & -0.097 & -0.490 & -0.277 & 0.277 & 0.490 & 0.097 & -0.415 \\ 0.353 & -0.353 & -0.353 & 0.353 & 0.353 & -0.353 & -0.353 & 0.353 \\ 0.277 & -0.490 & 0.097 & 0.415 & -0.415 & -0.097 & 0.490 & -0.277 \\ 0.191 & -0.461 & 0.461 & -0.191 & -0.191 & 0.461 & -0.461 & 0.191 \\ 0.097 & -0.277 & 0.415 & -0.490 & 0.490 & -0.415 & 0.277 & -0.095 \end{pmatrix}$$

#### 4. 4 2次元DCTの効果

- 2次元DCTで、画像は空間領域から(一種の)周波数領域に変換される。変換後のDCT値の絶対値は、直流成分が非常に大きな値になる(次ページ)。交流成分は直流成分よりもかなり小さく、周波数が高くなるにつれて小さくなる傾向がある。
- 写真画像がこの性質を持つことを利用してデータ量を圧縮する。すなわち、直流成分に多くのビット数を割り当て、交流成分に少ないビット数を割り当てる。交流成分の中では、周波数が高くなるにつれてビット数を少なくする。このように、絶対値が大きくなる成分に多くのビットを割り当て、絶対値が小さくなる成分に少ないビットを割り当てることで、全体のデータ量を削減することが可能になる。

#### 4. 4 2次元DCTの効果

#### 8×8小領域の画素値の例

$$A = \begin{pmatrix} 167 & 173 & 156 & 132 & 163 & 173 & 170 & 170 \\ 167 & 174 & 153 & 125 & 159 & 170 & 167 & 169 \\ 168 & 174 & 152 & 133 & 168 & 172 & 166 & 168 \\ 169 & 172 & 147 & 135 & 168 & 171 & 167 & 167 \\ 170 & 173 & 147 & 136 & 167 & 170 & 170 & 171 \\ 169 & 172 & 140 & 136 & 171 & 171 & 169 & 167 \\ 170 & 174 & 144 & 142 & 173 & 171 & 168 & 169 \\ 173 & 175 & 137 & 135 & 170 & 171 & 170 \end{pmatrix}$$

交流成分は、周波数が高くなるにつれて、 値が小さくなる

#### 直流成分は大きい値になる

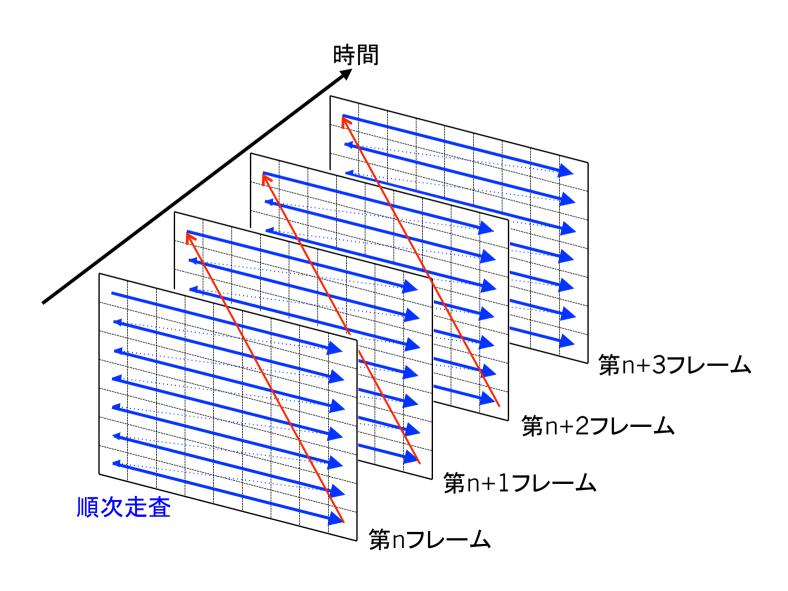
$$dct = \begin{bmatrix} 1299.50 & -21.87 & 60.31 & 58.40 & -20.25 & -52.35 & -11.18 & 13.72 \\ -3.14 & 2.58 & 3.12 & -9.67 & -13.93 & -0.01 & 11.79 & 5.45 \\ 0.75 & -0.49 & 3.30 & -0.23 & -2.54 & -0.51 & 0.47 & -0.43 \\ 1.12 & -2.13 & -1.29 & -1.65 & -0.41 & 2.50 & -0.30 & 0.07 \\ 1.75 & -1.30 & 2.27 & 0.30 & 0.50 & 2.02 & -0.40 & -0.52 \\ 4.01 & 0.84 & -7.56 & -0.39 & 2.76 & 0.88 & -0.20 & -0.65 \\ 0.50 & -1.17 & -1.53 & 3.30 & 1.05 & -0.72 & -1.30 & 0.03 \\ 5.02 & 0.56 & -1.72 & -3.46 & 0.52 & 0.42 & 0.97 & 0.18 \end{bmatrix}$$

8×8小領域の画素値を2次元DCTした値

#### 4.5 動画像

- 動画像(動画、映像)は、一定の時間間隔でサンプリングした静止画像の時系列である。
- 動画像を構成する一枚の画像をフレーム、毎秒あたりのフレーム数をフレームレートと呼ぶ。
- 動画像は水平軸、垂直軸、時間軸を有する3次元信号であるが、 伝送あるいはコンピュータ処理する場合、次ページのような走査 (スキャン)によって、1次元化することが一般的である。
- 次ページのものを順次走査と呼ぶ。これ以外に、フレーム内の走査線を奇数と偶数に分け、奇数ラインだけの画像をスキャンしたのち、偶数ラインだけの画像をスキャンする飛び越し走査がある。奇数ラインだけの画像および偶数ラインだけの画像をフィールドと呼ぶ。

#### 4.5 動画像



#### 4. 6 映像(動画)圧縮技術

- 広く利用されている映像圧縮技術にMPEG-2とH.264/
  MPEG-4 AVC、モーションJPEGなどがある。モーションJPEGは、
  各フレームを静止画として独立にJPEG圧縮するものである。
- MPEG-2はDVD Video、国内の地上波デジタル放送で利用されている。H.264/MPEG-4はBL(ブルーレイ)、その他のデジタル放送で利用されている。デジタルカメラで撮影する動画やインターネットで流通する動画でも利用されている。
- HD動画をデジタル伝送するためには映像圧縮技術は不可欠である。非圧縮と比較して、MPEG-2で40分の1程度、H.264で8 0分の1程度に圧縮される。
- MPEG映像圧縮技術は、静止画圧縮技術と「動き補償予測を用いたフレーム間符号化法」のハイブリット符号化方式である。

#### 4. 6 映像(動画)圧縮技術

圧縮方式	伝送レート	記録時間(16GB)
非圧縮	1Gbps	32秒
MPEG-2	24Mbps	1時間30分
H.264/MPEG-4 AVC	12Mbps	3時間

HD動画を伝送する場合の圧縮方式の比較

#### 4.7 MPEG符号化のデータ構造

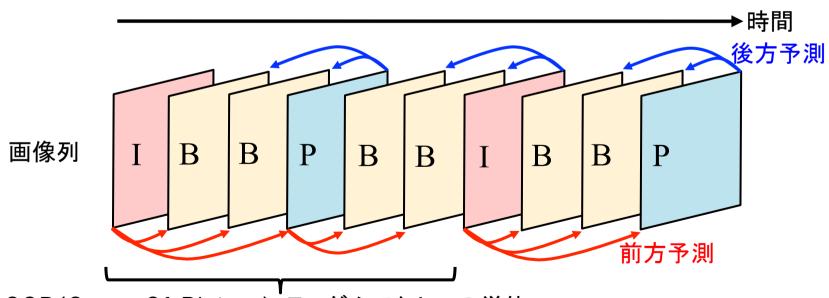
- 圧縮後の動画像のフレームはIピクチャー(intra-coded picture)、Pピクチャー(predictive-coded picture)、Bピクチャー(bidirectionally predictive-coded picture)のどれかーつに割り当てられる。
- Ⅰピクチャーはフレーム内符号化されるフレーム
- Pピクチャーは順方向動き補償予測を用いてフレーム間符号化 されるフレーム
- Bピクチャーは双方向動き補償予測を用いてフレーム間符号化 されるフレーム
- GOP(Group Of Picture)は、1枚以上のIピクチャーを含むフレームの集まりである。Iピクチャーを用いることでランダムアクセスや高速再生などを実現することができる。

#### 4.7 MPEG符号化のデータ構造

Ⅰピクチャー:フレーム間予測なし、圧縮率△

Pピクチャー: フレーム間予測あり(前方予測のみ)、圧縮率O

Bピクチャー:フレーム間予測あり(前方・後方・双方向)、圧縮率◎

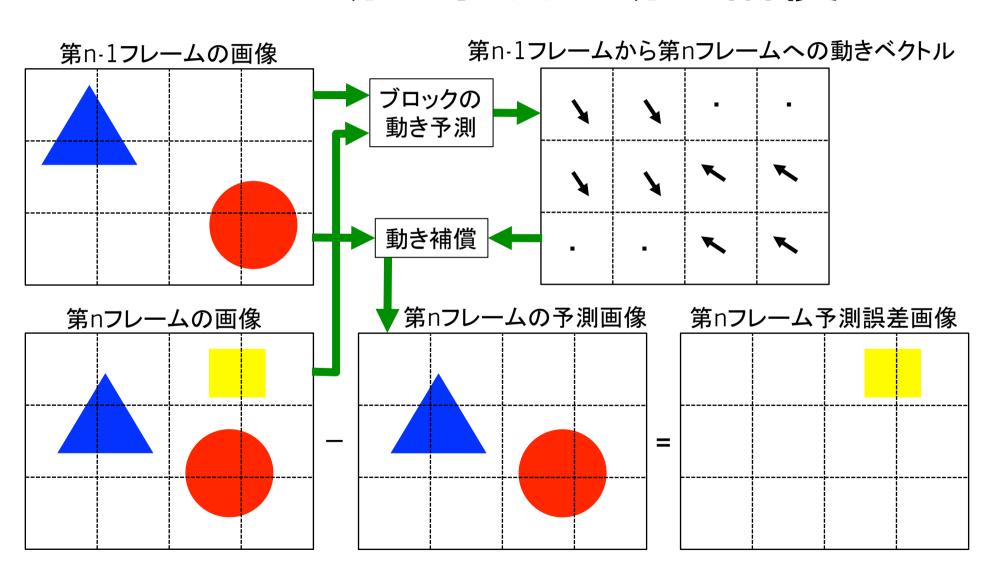


GOP(Group Of Picture):ランダムアクセスの単位

#### 4.8 動き予測と動き補償

- MPEG圧縮の重要部分は、動き予測に基づくフレーム間符号化である。これを、次ページのように第n-1フレームから第nフレームに画像が変化した場合を例に説明する。
- フレームは固定サイズのブロックに分割され、ブロック単位で処理される。連続するフレームの対応するブロック近傍を比較することでブロック単位の動きベクトルを求める。この処理を動き予測と呼ぶ。
- 第n-1フレームに、計算された動きベクトルを適用することで、第 nフレームの予測画像を生成する。入力画像と予測画像の差分 をとることで予測誤差画像を生成する。この予測誤差画像は、元 の入力画像と比較して信号変化が少ない画像になるため、これ を符号化することでデータ量を圧縮することが可能になる。

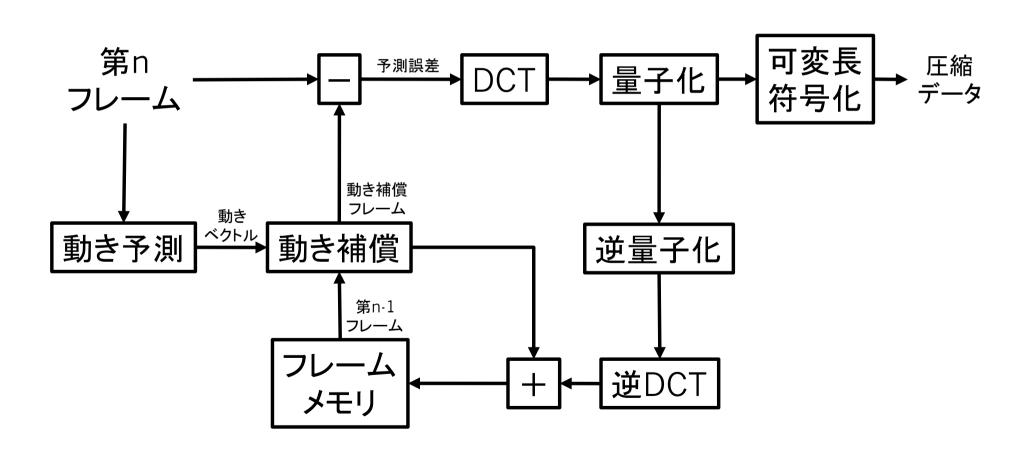
#### 4.8 動き予測と動き補償



#### 4.9 MPEG符号化の処理フロー

- MPEG符号化の処理フローを次ページに示す。第nフレームが入力された段階で、フレームメモリに第n-1フレームが保存されている。
- 第n-1フレームと第nフレームから動き予測と動き補償を行うことで、動き補償フレームが計算される。
- 動き補償フレームと第nフレームの差分を2次元DCTする。 DCT値を量子化した後、可変長符号化することで圧縮され たデータを得る。
- 一方、差分画像を空間領域に戻した後、動き補償フレームと加算することで第nフレームを再構築する。再構築された第nフレームは、第n+1フレームの符号化に利用される。

#### 4.9 MPEG符号化の処理フロー



#### 参考になる書籍・文献

- 「デジカメの画像処理」、蚊野浩監修/映像情報メディア学会編、オーム社、平成2 3年
- 「ディジタル画像処理[改訂新版]」、ディジタル画像処理編集委員会、CG-ARTS協会、2015年
- 「画像情報符号化」、映像情報メディア学会編、コロナ社、2008年
- 「詳解OpenCV コンピュータビジョンライブラリを使った画像処理・認識」、Gary Bradski, Adrian Kaebler著、松田晃一訳、オライリー・ジャパン、2009年
- 「CCD・CMOSカメラ技術入門」、竹村裕夫著、コロナ社、2008年
- 「画像入力とカメラ」、寺西信一監修/電子情報通信学会編、オーム社、平成24年
- 「CMOSイメージセンサ」、相澤清晴・浜本隆之編著、コロナ社、2012年
- 「講座:画像入力デバイスの基礎[全6回]」、映像メディア学会誌、Vol.68、No.1~No.6、2014年

#### 代表的な静止画像ファイル形式

名称	拡張子	説明・特徴
JPEG	.jpg	通常は非可逆圧縮で用い,写真画像に適した形式.デジカメでは1/5程度の圧縮率で利用される.
JPEG 2000	.jp2など	JPEGの課題である高圧縮率におけるモスキートノイズ,ブロックノイズを改善した方式. JPEGとの互換性は無い.
TIFF	.tif	さまざまな種類の画像データに対応するメタフォーマット. 非圧縮で用いられることが多い. ソフトウェア間の互換ファイルとして利用されることが多い.
PNG	.png	可逆圧縮で画質の劣化がない. Webページで標準的に利用される.
GIF	.gif	PNG以前にインターネット上で標準的に利用されていた形式. 256色までの色数. アニメーション機能を含む.
ВМР	.bmp	Windowsにおける標準フォーマット. 通常, 非圧縮.
PICT	.pct	Macにおける標準フォーマット.
PostScript	.ps	文字や図形、画像を含む印刷物のフォーマットをプログラムする言語、及びこの言語で記述されたファイル、JPEG、PNGなどを扱うことができ、その静止画が印刷される位置、大きさを指定する。
PDF	.pdf	PostScriptと同様の特性を持ち、印刷に限らず、電子文書の標準的なファイル形式

#### 代表的な動画像ファイル形式

名称	拡張子	説明▪特徴
AVI	.avi	Windowsで動画を取り扱うときの標準的なフォーマット. さまざまなデータ形式(圧縮/非圧縮)を利用できるメタフォーマット.
Windows Media	.wmv	Windows Media Playerで再生可能な動画コーデックとして開発されたもの.
Quick Time	.mov	Macで動画を取り扱うときの標準的なフォーマット. さまざまなデータ形式(圧縮/非圧縮)を利用できるメタフォーマット.
MPEG2	.mpg	MPEG1~MPEG4はいずれも動画コーデックの名称である.その中で,MPEG2はDVDビデオや地上波デジタル放送で広く用いられている.
MPEG4/AVC	.mp4	MPEG4には多くの規格が含まれるが,その中でMPEG4/AVCはH.264ともよばれ,現在,最も広く利用されている動画コーデックである.圧縮率はMPEG2の2倍程度である.
Motion JPEG	.mjp	JPEG圧縮された静止画が順次格納された形式.
Animation GIF	.gif	複数のGIF画像が順次格納された形式.