High resolution image fusion via fusion frames

Shidong Li

San Francisco State University

jointly with
Zhenjie Yao and *Weidong Yi*
Graduate University of the Chinese Academy of Sciences

FFT - University of Maryland

February 17 & 18, 2011

Outline

Outline

What's a Frame?

Definition of frames

A sequence $\{x_n\}_{n \in I}$ in $\mathcal{H}$ is a ***frame*** for $\mathcal{H}$, if there exist $0 < A \leq B < \infty$ (*lower* and *upper frame bounds*) such that

$$A\|f\|^2 \leq \sum_{n \in I} |\langle f, x_n \rangle|^2 \leq B\|f\|^2 \quad \text{f or all } f \in \mathcal{H}.$$

What's a Frame?

Definition of frames

A sequence $\{x_n\}_{n\in I}$ in $\mathcal{H}$ is a *frame* for $\mathcal{H}$, if there exist $0 < A \le B < \infty$ (*lower* and *upper frame bounds*) such that

$$A\|f\|^2 \le \sum_{n\in I} |\langle f, x_n\rangle|^2 \le B\|f\|^2 \quad \text{for all } f \in \mathcal{H}.$$

Basic properties of frames

- Frame is a "basis-like" system in $\mathcal{H}$.
- Frame is typically redundant.
- Frame representation: there are *dual frames* $\{\tilde{x}_n\} \subseteq \mathcal{H}$ such that

$$\text{for all } f \in \mathcal{H}, \quad f = \sum_n \langle f, \tilde{x}_n\rangle x_n = \sum_n \langle f, x_n\rangle \tilde{x}_n.$$

Why Fusion Frames?

- General Problem: perform data fusion from measurements out of a sensor network or distributed systems.

Why Fusion Frames?

- General Problem: perform data fusion from measurements out of a sensor network or distributed systems.
- Each sensor S_i has its spanning subspace W_i.

Why Fusion Frames?

- General Problem: perform data fusion from measurements out of a sensor network or distributed systems.
- Each sensor S_i has its spanning subspace W_i.
- Subspaces $\{W_i\}$ can have arbitrary overlap - *much like overlaps in frames*.

Why Fusion Frames?

- General Problem: perform data fusion from measurements out of a sensor network or distributed systems.
- Each sensor S_i has its spanning subspace W_i.
- Subspaces $\{W_i\}$ can have arbitrary overlap - *much like overlaps in frames*.
- Sensor subspaces are almost never orthogonal - *much like the non-orthogonality in frames*.

Why Fusion Frames?

- General Problem: perform data fusion from measurements out of a sensor network or distributed systems.
- Each sensor S_i has its spanning subspace W_i.
- Subspaces $\{W_i\}$ can have arbitrary overlap - *much like overlaps in frames*.
- Sensor subspaces are almost never orthogonal - *much like the non-orthogonality in frames*.
- Sensor network is almost always redundant - *much like the redundancy in frames*.

Why Fusion Frames?

- General Problem: perform data fusion from measurements out of a sensor network or distributed systems.
- Each sensor S_i has its spanning subspace W_i.
- Subspaces $\{W_i\}$ can have arbitrary overlap - *much like overlaps in frames*.
- Sensor subspaces are almost never orthogonal - *much like the non-orthogonality in frames*.
- Sensor network is almost always redundant - *much like the redundancy in frames*.
- Need a technique to perform data fusion among a set of overlapping, non-orthogonal and redundant data measurements, regardless how complicated the sensor subspaces $\{W_i\}$ are related.

Outline

Definition of Fusion Frames

Fusion Frames

Let $\{W_i\}_{i\in I}$ be a set of closed subspaces of a Hilbert space $\mathcal{H}$, and let $\{\pi_{W_i}\}$ be orthogonal projections onto $\{W_i\}$, and let $\{v_i > 0\}$ be weights. $\{\pi_{W_i}, v_i\}$ is a *fusion frame* for $\mathcal{H}$, if there exist constants $0 < C \leq D < \infty$ such that

$$C\|f\|^2 \leq \sum_{i\in I} v_i^2 \|\pi_{W_i}(f)\|^2 \leq D\|f\|^2 \quad \text{for all } f \in \mathcal{H}. \qquad (1)$$

Definition of Fusion Frames

Fusion Frames

Let $\{W_i\}_{i\in I}$ be a set of closed subspaces of a Hilbert space $\mathcal{H}$, and let $\{\pi_{W_i}\}$ be orthogonal projections onto $\{W_i\}$, and let $\{v_i > 0\}$ be weights. $\{\pi_{W_i}, v_i\}$ is a *fusion frame* for $\mathcal{H}$, if there exist constants $0 < C \leq D < \infty$ such that

$$C\|f\|^2 \leq \sum_{i\in I} v_i^2 \|\pi_{W_i}(f)\|^2 \leq D\|f\|^2 \quad \text{for all } f \in \mathcal{H}. \tag{1}$$

[CK'04] P. Casazza, G. Kutyniok, Frames of Subspaces, *Contemp. Math.*, Amer. Math. Soc., **345** (87-113), 2004.

[CKL'06] P. Casazza, G. Kutyniok and S. Li, Fusion Frames and Distributed Systems, *Appl. Comp. Harmon. Anal.*, **25** (114 - 132), 2008.

Examples of related work on fusion frames

- **[Sun'06]** W. Sun, G-frames and g-Riesz bases, *J. of Math. Anal. and Appl.*, **322**:1, (437 - 452), 2006.
- **[Kaftal, Larson and Zhang, 2009]** Operator Valued Frames
- Others...

On related work

- While g-frame or operator-valued frames are indeed more general, there is a reason why projection operators are particularly important in applications.

On related work

- While g-frame or operator-valued frames are indeed more general, there is a reason why projection operators are particularly important in applications.
- Because physical devices are naturally modeled by a frame, and thereby a natural projection operator.

On related work

- While g-frame or operator-valued frames are indeed more general, there is a reason why projection operators are particularly important in applications.
- Because physical devices are naturally modeled by a frame, and thereby a natural projection operator.
- The restriction to projection operators is a stronger constraint. The analysis of projection-based fusion frames becomes different and a bit more involved.

The Fusion Frame Operator S_W

- *Analysis operator* $T_{\mathcal{W}} : \mathcal{H} \to \left(\{W_i\} \right)_{\ell_2}$

$$T_{\mathcal{W}}(f) = \{v_i \pi_{W_i}(f)\}_{i \in I}.$$

Then the *fusion frame operator* $S_{\mathcal{W}}$ is defined by

$$S_{\mathcal{W}}(f) = T_{\mathcal{W}}^* T_{\mathcal{W}}(f) = \sum_{i \in I} v_i^2 \pi_{W_i}(f).$$

The Fusion Frame Operator S_W

- *Analysis operator* $T_{\mathcal{W}} : \mathcal{H} \to \left(\{W_i\}\right)_{\ell_2}$

$$T_{\mathcal{W}}(f) = \{v_i \pi_{W_i}(f)\}_{i\in I}.$$

Then the *fusion frame operator* $S_{\mathcal{W}}$ is defined by

$$S_{\mathcal{W}}(f) = T_{\mathcal{W}}^* T_{\mathcal{W}}(f) = \sum_{i\in I} v_i^2 \pi_{W_i}(f).$$

- Fusion frame inequality is equivalent to

$$CI_d \leq S_{\mathcal{W}} \leq DI_d.$$

A computational difference

Theorem: ***Global processing***

Let $\{x_{i,j}\}_{j\in J_i}$ be a frame for W_i. $\{\pi_{W_i}, v_i\}$ is a fusion frame of $\mathcal{H}$ iff $\{v_i x_{i,j}\}_{i,j}$ is a frame of $\mathcal{H}$ (with the frame operator S_F).

$$\forall f \in \mathcal{H}, \quad f = \sum_{i,j} \langle f, v_i x_{i,j} \rangle S_F^{-1} v_i x_{i,j}.$$

A computational difference

Theorem: ***Global processing***

Let $\{x_{i,j}\}_{j\in J_i}$ be a frame for W_i. $\{\pi_{W_i}, v_i\}$ is a fusion frame of $\mathcal{H}$ iff $\{v_i x_{i,j}\}_{i,j}$ is a frame of $\mathcal{H}$ (with the frame operator S_F).

$$\forall f \in \mathcal{H}, \quad f = \sum_{i,j} \langle f, v_i x_{i,j}\rangle S_F^{-1} v_i x_{i,j}.$$

Theorem: ***Parallel and local processing***

Let $\{x_{i,j}\}_{j\in J_i}$ be a frame for W_i with a dual frame $\{\tilde{x}_{i,j}\}_{j\in J_i}$. Then $\forall f \in \mathcal{H}, \quad \pi_{W_i} f = \sum_{j\in J_i} \langle f, x_{i,j}\rangle \tilde{x}_{i,j}$. Consequently, for all $f \in \mathcal{H}$,

$$f = S_W^{-1}\left(\sum_i v_i^2 \pi_{W_i} f\right) = \sum_i v_i^2 \sum_j \langle f, x_{i,j}\rangle S_W^{-1}\left(\tilde{x}_{i,j}\right). \tag{2}$$

Outline

Image fusion formulation

- Every observed pixel $g(m,n)$ of a (discrete) image corresponds to a bounded linear functional acting on image f.

Image fusion formulation

- Every observed pixel $g(m,n)$ of a (discrete) image corresponds to a bounded linear functional acting on image f.
- By the Riesz Representation Theorem, there is a function $h_{m,n}$ such that $g(m,n) = \langle f, h_{m,n} \rangle$.

Image fusion formulation

- Every observed pixel $g(m, n)$ of a (discrete) image corresponds to a bounded linear functional acting on image f.
- By the Riesz Representation Theorem, there is a function $h_{m,n}$ such that $g(m, n) = \langle f, h_{m,n} \rangle$.
- We actually know exactly what $\{h_{m,n}\}_{m,n}$ is.

Image fusion formulation

- Every observed pixel $g(m, n)$ of a (discrete) image corresponds to a bounded linear functional acting on image f.
- By the Riesz Representation Theorem, there is a function $h_{m,n}$ such that $g(m, n) = \langle f, h_{m,n} \rangle$.
- We actually know exactly what $\{h_{m,n}\}_{m,n}$ is.
- Let $r_i(x, y)$ ($i \in I$) be the *Impulse Response Function* (IRF) of the i^{th} camera. Let $f(x, y)$ be the image source function.

Image fusion formulation

- Every observed pixel $g(m,n)$ of a (discrete) image corresponds to a bounded linear functional acting on image f.
- By the Riesz Representation Theorem, there is a function $h_{m,n}$ such that $g(m,n) = \langle f, h_{m,n} \rangle$.
- We actually know exactly what $\{h_{m,n}\}_{m,n}$ is.
- Let $r_i(x,y)$ ($i \in I$) be the *Impulse Response Function* (IRF) of the i^{th} camera. Let $f(x,y)$ be the image source function.
- Then the image $g_i(x,y)$ is given by

$$g_i(x,y) = (f * r_i)(x,y) = \langle f, h_i(\cdot - x, \cdot - y) \rangle, \tag{3}$$

where $h_i(x,y) \equiv \overline{r_i(-x,-y)}$.

Image fusion formulation (cont'd)

- Suppose observation sample points are $x \equiv \{x_m\}_m$ and $y = \{y_n\}_n$.

Image fusion formulation (cont'd)

- Suppose observation sample points are $x \equiv \{x_m\}_m$ and $y = \{y_n\}_n$.
- Then the observed (discrete) image $\{g_i(x_m, y_n)\}_{m,n}$ is given by

$$g_i(x_m, y_n) = \langle f, h_i(\cdot - x_m, \cdot - y_n) \rangle. \tag{4}$$

Image fusion formulation (cont'd)

- Suppose observation sample points are $x \equiv \{x_m\}_m$ and $y = \{y_n\}_n$.
- Then the observed (discrete) image $\{g_i(x_m, y_n)\}_{m,n}$ is given by
$$g_i(x_m, y_n) = \langle f, h_i(\cdot - x_m, \cdot - y_n)\rangle. \quad (4)$$
- That is, the observed image $\{g_i(x_m, y_n)\}$ is the set of transformation coefficients of the source function f with respect to the ***camera frame*** $\{h_i(\cdot - x_m, \cdot - y_n)\}_{m,n}$ of the i^{th} camera subspace W_i.

Image fusion formulation (cont'd)

- Consequently, multi-camera image fusion is ***extremely naturally*** the fusion frame problem on the set of camera subspaces $\{W_i\}$ in $\mathcal{H} \equiv \sum_i W_i$. Namely, to combine $f_i \in W_i$ to obtain a function $f \in \mathcal{H}$.

Image fusion formulation (cont'd)

- Consequently, multi-camera image fusion is ***extremely naturally*** the fusion frame problem on the set of camera subspaces $\{W_i\}$ in $\mathcal{H} \equiv \sum_i W_i$. Namely, to combine $f_i \in W_i$ to obtain a function $f \in \mathcal{H}$.

Camera functions as a projection operator

Let $\{\tilde{h}_{i;m,n}\}_{m,n}$ be a dual frame of $\{h_{i;m,n}\}_{m,n}$ (the frame spanning the (observation) space W_i of the camera). Then the true image observed is an orthogonal projection of the image f onto W_i:

$$\pi_{W_i} f = \sum_{m,n} \langle f, h_{i;m,n} \rangle \tilde{h}_{i;m,n}.$$

Image fusion formulation (cont'd)

- Consequently, multi-camera image fusion is ***extremely naturally*** the fusion frame problem on the set of camera subspaces $\{W_i\}$ in $\mathcal{H} \equiv \sum_i W_i$. Namely, to combine $f_i \in W_i$ to obtain a function $f \in \mathcal{H}$.

Camera functions as a projection operator

Let $\{\tilde{h}_{i;m,n}\}_{m,n}$ be a dual frame of $\{h_{i;m,n}\}_{m,n}$ (the frame spanning the (observation) space W_i of the camera). Then the true image observed is an orthogonal projection of the image f onto W_i:

$$\pi_{W_i} f = \sum_{m,n} \langle f, h_{i;m,n} \rangle \tilde{h}_{i;m,n}.$$

- This is why typical sensor measurement functions as a projection operator, which is why projection operators are the subject of study in fusion frame!

Image fusion formulation (cont'd)

Fusion through the fusion frame operator

Recall that the *fusion frame operator* $S_W : \mathcal{H} \to \mathcal{H}$

$$\forall f \in \mathcal{H}, \quad S_W f = \sum_i v_i^2 \pi_{W_i} f.$$

Image fusion formulation (cont'd)

Fusion through the fusion frame operator

Recall that the *fusion frame operator* $S_W : \mathcal{H} \to \mathcal{H}$

$$\forall f \in \mathcal{H}, \quad S_W f = \sum_i v_i^2 \pi_{W_i} f.$$

How to fuse together multiple observations?

By simply applying S_W^{-1}:

$$\begin{aligned} \forall f \in \mathcal{H}, \quad f &= S_W^{-1} \left(\sum_i v_i^2 \pi_{W_i} f \right) \\ &= \sum_i v_i^2 \sum_{m,n} \langle f, h_{i;m,n} \rangle S_W^{-1} \left(\tilde{h}_{i;m,n} \right). \end{aligned} \tag{5}$$

An equivalent fusion formulation

Equivalent formulation: *composite camera frame*

Recall, $\{\pi_{W_i}, v_i\}$ is a fusion frame of $\mathcal{H}$ *iff* $\{v_i h_{i;m,n}\}_{i;m,n}$ is a frame of $\mathcal{H}$. We may term $\{v_i h_{i;m,n}\}_{i;m,n}$ the *composite camera frame* with the composite frame operator S_F.

An equivalent fusion formulation (cont'd)

Fusion through *composite camera frame* (cont'd)

Therefore, fused image f may also be given by

$$f = \sum_{i;m,n} \langle f, v_i h_{i;m,n} \rangle S_F^{-1} v_i h_{i;m,n}.$$

or (by the fusion frame approach, for comparison)

$$f = \sum_{i} \sum_{m,n} \langle f, v_i h_{i;m,n} \rangle S_W^{-1} \left(v_i \tilde{h}_{i;m,n} \right)$$

Remark: $S_F^{-1} h_{i;m,n} \neq S_W^{-1} \left(v_i \tilde{h}_{i;m,n} \right)$, in general!

Outline

Dual evaluation: A *Dimension Invariance Principle*

- Let $\{T_{kc}g^{(j)}\}_{j,k}$ and $\{T_{kc}h^{(j)}\}_{j,k}$ be a pair of dual frame sequences of translates. Define $|\text{Supp}\{g^{(j)}\}|$ the measure of the smallest "interval" containing the support of all the $g^{(j)}$'s.

Theorem [Jointly with J. Cahill]

Let $\mathcal{H} = \mathbb{F}^n$. Suppose $\{T_{kc}g^{(j)}\}$ and $\{T_{kc}h^{(j)}\}$ are a dual pair of frames of translates for $\mathcal{X} = \text{span}\{T_{kc}g^{(j)}\} \subseteq \mathcal{H}$ which also satisfies

$$2\left|\text{Supp}\{g^{(j)}\}\right| + \left|\text{Supp}\{h^{(j)}\}\right| \leq n. \qquad (6)$$

Let $\tilde{\mathcal{H}} = \mathbb{F}^{\tilde{n}}$ where c divides $\tilde{n} \geq n$, then $\{T_{kc}\tilde{g}^{(j)}\}$ and $\{T_{kc}\tilde{h}^{(j)}\}$ remain a dual pair of frames of translates for $\tilde{\mathcal{X}} = \text{span}\{T_{kc}\tilde{g}^{(j)}\} \subseteq \tilde{\mathcal{H}}$. Here $\tilde{g} \in \mathbb{F}^{\tilde{n}}$ is the natural embedding of $g \in \mathbb{F}^n$.

Calculating dual frames (cont'd)

Remarks:

The *Dimension Invariance Principle* applies to a number of scenarios.

- Non-uniform but proportional translates;
- Uniform and proportional non-uniform multi-Gabor frames;
- Uniform and proportional non-uniform multi-variable frames of translates, etc.

Calculating dual frames (cont'd)

Theorem

Approximate duals through truncation is also stable. Suppose that $\|\tilde{h}^{(j)} - \tilde{h}_a^{(j)}\| \leq \delta$. Then

$$||f - \sum_{j,k} \langle f, T_{kc}\tilde{h}_a^{(j)} \rangle T_{kc}\tilde{g}^{(j)}||_1 \leq C_1\delta||f||_1$$

and

$$||f - \sum_{j,k} \langle f, T_{kc}\tilde{g}^{(j)} \rangle T_{kc}\tilde{h}_a^{(j)}||_1 \leq C_1\delta||f||_1$$

Iterative Algorithms

Iterative Algorithms

- Ultimately, observation equations are given by $Hf = g$. Here H is a circulant matrix formed by the translates of the PSF/IRF of the camera/sensor.

Iterative Algorithms

- Ultimately, observation equations are given by $Hf = g$. Here H is a circulant matrix formed by the translates of the PSF/IRF of the camera/sensor.

Algorithm 1

Let $\tilde{H}$ be a low-pass operator satisfying

$$\tilde{H}H + R = I, \tag{7}$$

Assume that $\tilde{H}$ is such that $0 \leq \tilde{H}H \leq I$. Let $f_0 = \mathbf{0}$. Define

$$f_{n+1} = \tilde{H}Hf + Rf_n = \tilde{H}g + Rf_n.$$

Then the sequence of images $\{f_n\} \to f$ in the Euclidean norm.

Iterative Algorithms

- Ultimately, observation equations are given by $Hf = g$. Here H is a circulant matrix formed by the translates of the PSF/IRF of the camera/sensor.

Algorithm 1

Let $\tilde{H}$ be a low-pass operator satisfying

$$\tilde{H}H + R = I, \tag{7}$$

Assume that $\tilde{H}$ is such that $0 \leq \tilde{H}H \leq I$. Let $f_0 = \mathbf{0}$. Define

$$f_{n+1} = \tilde{H}Hf + Rf_n = \tilde{H}g + Rf_n.$$

Then the sequence of images $\{f_n\} \to f$ in the Euclidean norm.

Remark: The choice of $\tilde{H}$ can be rather flexible. We tried $\tilde{H} = (H')^m$, and observed that higher m is good for low SNR.

Iteration Algorithm 2

Algorithm 2

A direct generalization to *Algorithm 1* would be to introduce a multiplicative factor $(1-\beta)$ (with a small $0 < \beta << 1$) in front of Rf_n components so as to suppress high-frequency noises in the iteration:

$$f_{n+1} = \tilde{H}g + (1-\beta)Rf_n. \tag{8}$$

Iteration Algorithm 3

Algorithm 3

Another generalization to *Algorithm 1* would be to apply a soft thresholding to the term Rf_n in *Algorithm 1*, and results in thresholding de-noising effects out of high-frequency components.

$$f_{n+1} = \tilde{H}g + \mathcal{T}\left(Rf_n\right), \tag{9}$$

where $\mathcal{T}$ is some (soft) thresholding operator.

Fusion system calibration and alignment

Fusion system calibration and alignment

Three basic steps

1. Dynamic programming for imaging intensity calibration.

Fusion system calibration and alignment

Three basic steps

1. Dynamic programming for imaging intensity calibration.
2. Affine transoformation for calibrations of image scale, rotation and translation variations.

Fusion system calibration and alignment

Three basic steps

1. Dynamic programming for imaging intensity calibration.
2. Affine transoformation for calibrations of image scale, rotation and translation variations.
3. (Sub-pixel) alignments.

Outline

Simulating 4 pictures by one camera

Simulating 4 pictures by one camera

(c) One observed LR image, $\sigma = 0.7$

(d) $\sigma = \sigma_e = [0.7, 0.7, 0.7, 0.7]$, interleaved coarse image

Simulating 4 pictures by one camera (cont'd)

(e) Original image

(f) $\sigma = \sigma_e$, dimension invariance

Figure: $\sigma_e = [0.7, 0.7, 0.7, 0.7]$

Simulating 4 pictures by one camera (cont'd)

(a) $\sigma = \sigma_s$, dimension invariance (b) $\sigma = \sigma_b$, dimension invariance

Figure: $\sigma_s = [0.6, 1, 1.2, 0.8]$, $\sigma_b = [0.7, 2, 0.9, 1.5]$

Side-by-side comparison

(a) $\sigma = \sigma_e$, Dimension Invariance (b) $\sigma = \sigma_b$, Dimension Invariance

Figure: Fusion results comparison between correct PSF estimation, and incorrect PSF estimation

Simulating 4 images by 4 different cameras

Simulating 4 images by 4 different cameras

(e) LR image 1, $\sigma(1) = 0.7$

(f) LR image 2, $\sigma(2) = 2$

(g) LR image 3, $\sigma(3) = 0.9$

(h) LR image 4, $\sigma(4) = 1.5$

Simulating 4 images by 4 different cameras (cont'd)

(i) Fused image, inaccurate PSF estimation

(j) Fused image, accurate PSF estimation

HR image fusion: 4 actual images by 1 camera at diff times

HR image fusion: 4 actual images by 1 camera at diff times

(o)

(p)

(q)

(r)

HR image fusion: 4 actual images by 1 camera at diff times (cont'd)

(s) After alignment

(t) Fused image

HR image fusion: 4 actual images taken by 4 different cameras

(u)

(v)

(w)

(x)

HR image fusion: 4 actual images taken by 4 different cameras (cont'd)

(y) After alignment

(z) Fused image

Thank You!