Constant Aspect-Ratio Tiling

Guillaume Iooss, Sanjay Rajopadhye, Christophe Alias, Yun Zou

Colorado State University - ENS Lyon

January 20, 2014

Parametric tiling

- **Tiling** is an important transformation:
  - $\rightarrow$ Locality improvement
  - $\rightarrow$ New level of granularity (can be exploited for parallelism)

  If the tile sizes are constant, polyhedral $\left(i = 4.\alpha + ii,\ 0 \leq ii < 4\right)$

- **Parametric tiling**: tiling where the tile sizes are parameters
  - $\rightarrow$ Tile size can be selected during runtime (autotuning)

  Not a polyhedral transformation $\left(i = b.\alpha + ii,\ 0 \leq ii < b\right)$

Parametric tiling - State of the art

- Parametric tiling is embedded in the code generation phase:
  - Fourier-Motzkin symbolic elimination [Gösslinger, CPC2004]
  - Tile the bounding box of the iteration domain [Lakshmi, PLDI2007]
  - *D-tiling* [Kim, LCPC10] (outset, inset)
  - *PrimeTile* [Hartono, ICS09], *DynTile* [Hartono, IDPDS10] and *PTile* [Baskaran, CGO10]
- Later transformations/analysis must be "hard-coded":
  - DynTile: find a wavefront schedule
  - Parametric GPU code generation [Athanasios, Kelly, LCPC13]

    → Exploit wavefront/rectangular parallelism

Contribution

Parametric tiling with *one tile size parameter* and fixed ratio for every dimensions → obtain a polyhedral program representation.

- *Example:* $b \times 2b$ and not $b \times c$

⇒ **Constant Aspect-Ratio Tiling** (CART)

Contribution

Parametric tiling with *one tile size parameter* and fixed ratio for every dimensions $\rightarrow$ obtain a polyhedral program representation.

- *Example:* $b \times 2b$ and not $b \times c$

$\Rightarrow$ **Constant Aspect-Ratio Tiling** (CART)

Rest of the talk: We focus on polyhedron and affine function

- Polyhedron $\xrightarrow{CART}$ union of "tiled" polyhedra
  - $\mapsto$ Improvment to have only one polyhedron per tiles
- Affine function $\xrightarrow{CART}$ piecewise affine function
  - In general, might admit modulo constraints
  - Under a condition, only admit polyhedral constraints

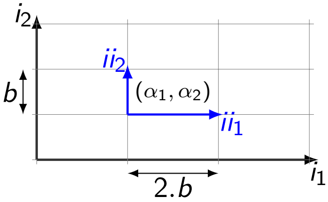
for $(i,j) \in \mathcal{D}$
    $A[i,j] = B[i+1, j-1]$

$A = \mathcal{D} : (i,j\text{->}i+1, j-1)@B$

Notations and CART hypothesis

Given a polyhedron $\mathcal{D}_{\vec{p}} = \{\vec{i} \mid \dots\}$:

- $\vec{i} = b.D.\vec{\alpha} + \vec{ii}$ where $\vec{0} \leq \vec{ii} < b.D.\vec{1}$
  - All dimensions tiled along canonical axis
  - $\vec{\alpha}/\vec{ii}$: blocked/local indices
  - b: tile size parameter
  - D: ratio (diagonal matrix)
- $\vec{p} = b.\vec{\lambda} + \vec{pp}$ where $\vec{0} \leq \vec{pp} < b.\vec{1}$
  - $\vec{\lambda}/\vec{pp}$: blocked/local parameters

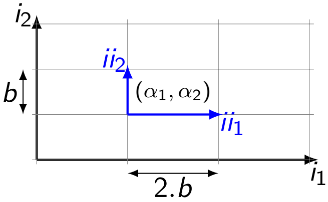


$$i_1 = 2.b.\alpha_1 + ii_1$$
$$i_2 = b.\alpha_2 + ii_2$$
$$D = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$

Notations and CART hypothesis

Given a polyhedron $\mathcal{D}_{\vec{p}} = \{\vec{i} \mid \dots\}$:

- $\vec{i} = b.D.\vec{\alpha} + \vec{ii}$ where $\vec{0} \leq \vec{ii} < b.D.\vec{1}$
  - All dimensions tiled along canonical axis
  - $\vec{\alpha}/\vec{ii}$: blocked/local indices
  - b: tile size parameter
  - D: ratio (diagonal matrix)
- $\vec{p} = b.\vec{\lambda} + \vec{pp}$ where $\vec{0} \leq \vec{pp} < b.\vec{1}$
  - $\vec{\lambda}/\vec{pp}$: blocked/local parameters



$$i_1 = 2.b.\alpha_1 + ii_1$$
$$i_2 = b.\alpha_2 + ii_2$$
$$D = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$

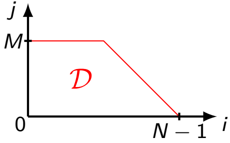
$\Rightarrow$ **Question:** How to obtain $\Delta_{\vec{\lambda},\vec{pp}} = \{\vec{\alpha}, \vec{ii} \mid \dots\}$?

Deriving the blocked union of polyhedra

- **Example:** $\mathcal{D} = \{i, j \mid i + j \leq N - 1 \ \wedge \ j \leq M \ \wedge \ 0 \leq i, j\}$ with tiles of size $b \times b$.

- Let us focus on the first constraint:

$$N - i - j - 1 \geq 0$$



Deriving the blocked union of polyhedra

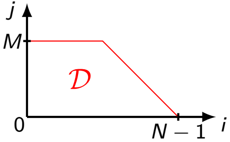
- **Example:** $\mathcal{D} = \{i, j \mid i + j \leq N - 1 \ \wedge \ j \leq M \ \wedge \ 0 \leq i, j\}$ with tiles of size $b \times b$.

- Let us focus on the first constraint:

$$N - i - j - 1 \geq 0$$

$$(N, i, j) = (N_{bl}, \alpha, \beta).b + (N_{loc}, ii, jj) \Updownarrow (substitution)$$

$$\boxed{(N_{bl} - \alpha - \beta).b} + (N_{loc} - ii - jj - 1) \geq 0$$



Deriving the blocked union of polyhedra

- **Example:** $\mathcal{D} = \{i,j \mid i+j \leq N-1 \ \wedge \ j \leq M \ \wedge \ 0 \leq i,j\}$ with tiles of size $b \times b$.
- Let us focus on the first constraint:

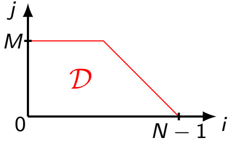
$$N - i - j - 1 \geq 0$$

$$\Updownarrow \ (substitution) \quad (N,i,j) = (N_{bl},\alpha,\beta).b + (N_{loc},ii,jj)$$

$$\boxed{(N_{bl} - \alpha - \beta).b} + (N_{loc} - ii - jj - 1) \geq 0$$

$$\Updownarrow \ b > 0$$

$$N_{bl} - \alpha - \beta + \boxed{\frac{N_{loc} - ii - jj - 1}{b}} \geq 0$$



Deriving the blocked union of polyhedra

- **Example:** $\mathcal{D} = \{i,j \mid i+j \leq N-1 \;\wedge\; j \leq M \;\wedge\; 0 \leq i,j\}$ with tiles of size $b \times b$.
- Let us focus on the first constraint:

$$N - i - j - 1 \geq 0$$

$$\Updownarrow \; (N,i,j) = (N_{bl},\alpha,\beta).b + (N_{loc},ii,jj) \;\text{(substitution)}$$

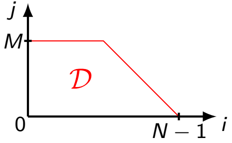
$$\boxed{(N_{bl} - \alpha - \beta).b} + (N_{loc} - ii - jj - 1) \geq 0$$

$$\Updownarrow \; b > 0$$

$$N_{bl} - \alpha - \beta + \boxed{\frac{N_{loc} - ii - jj - 1}{b}} \geq 0$$

$$\Updownarrow \; a \geq 0 \Leftrightarrow \lfloor a \rfloor \geq 0$$

$$N_{bl} - \alpha - \beta + \left\lfloor \frac{N_{loc} - ii - jj - 1}{b} \right\rfloor \geq 0$$



Computing k_{min} and k_{max}

- $k_1 = \left\lfloor \frac{N_{loc} - ii - jj - 1}{b} \right\rfloor$, where $0 \leq ii, jj, N_{loc} < b$, is bounded.
  - *Maximum of* k_1*:* reached for $N_{loc} = b - 1$, $ii = jj = 0$.

    $\rightarrow \ k_{1,max} = \left\lfloor \frac{(b-1)-1}{b} \right\rfloor = 0$
  - *Minimum of* k_1*:* reached for $N_{loc} = 0$, $ii = jj = b - 1$

    $\rightarrow \ k_{1,min} = \left\lfloor \frac{0-(b-1)-(b-1)-1}{b} \right\rfloor = -2 + \left\lfloor \frac{1}{b} \right\rfloor = -2$

$\Rightarrow k_1 \in \{-2, -1, 0\}$

Computing k_{min} and k_{max}

- $k_1 = \left\lfloor \frac{N_{loc} - ii - jj - 1}{b} \right\rfloor$, where $0 \leq ii, jj, N_{loc} < b$, is bounded.
  - *Maximum of k_1:* reached for $N_{loc} = b - 1$, $ii = jj = 0$.

    $\rightarrow \; k_{1,max} = \left\lfloor \frac{(b-1)-1}{b} \right\rfloor = 0$
  - *Minimum of k_1:* reached for $N_{loc} = 0$, $ii = jj = b - 1$

    $\rightarrow \; k_{1,min} = \left\lfloor \frac{0-(b-1)-(b-1)-1}{b} \right\rfloor = -2 + \left\lfloor \frac{1}{b} \right\rfloor = -2$

  $\Rightarrow k_1 \in \{-2, -1, 0\}$

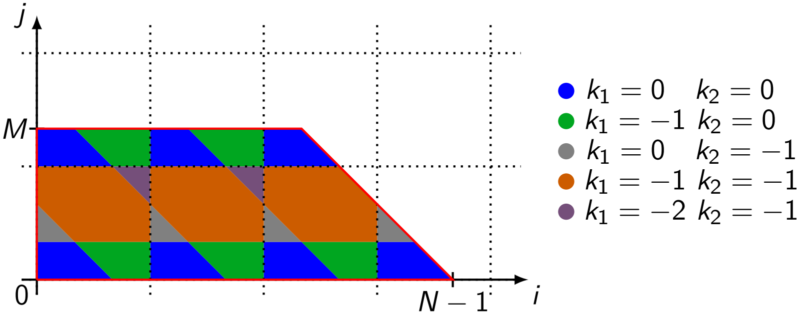
- Other constraints: $k_2 = \left\lfloor \frac{M_{loc} - jj}{b} \right\rfloor = -1 \text{ or } 0$, $k_3 = k_4 = 0$.

- We can derive the values of ii, jj corresponding to each k_i.

Deriving the blocked union of polyhedra: result

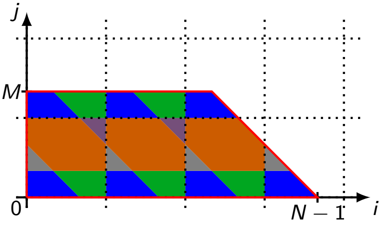
- Δ is the union of 6 polyhedra ($-2 \leq k_1 \leq 0$, $k_2 = -1$ or 0)

$$\Delta = \left\{ \alpha, \beta, ii, jj \;\middle|\; \begin{array}{c} N_{bl} - \alpha - \beta - 2 \geq 0 \\ M_{bl} - \beta - 1 \geq 0 \\ \alpha, \beta \geq 0 \\ -2.b \leq N_{loc} - ii - jj - 1 < -b \\ -b \leq M_{loc} - jj < 0 \\ 0 \leq ii, jj < b \\ (\text{for } k_1 = -2 \text{ and } k_2 = -1) \end{array} \right\} \cup \dots$$

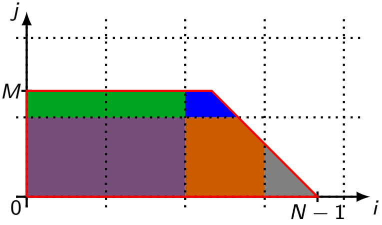


Merging tiles

We can merge them to obtain one polyhedron per tiles:



- $k_1 = 0 \quad k_2 = 0$
- $k_1 = -1 \ k_2 = 0$
- $k_1 = 0 \quad k_2 = -1$
- $k_1 = -1 \ k_2 = -1$
- $k_1 = -2 \ k_2 = -1$

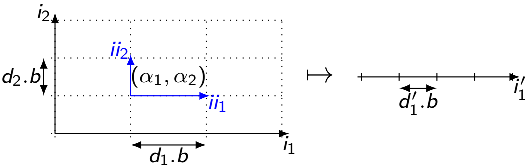


- $N_{bl} - \alpha - \beta = 0 \mid M_{bl} - \beta = 0$
- $N_{bl} - \alpha - \beta = 0 \mid M_{bl} - \beta \geq 1$
- $N_{bl} - \alpha - \beta = 1 \mid M_{bl} - \beta = 0$
- $N_{bl} - \alpha - \beta = 1 \mid M_{bl} - \beta \geq 1$
- $N_{bl} - \alpha - \beta \geq 2 \mid M_{bl} - \beta \geq 1$

Notations and CART hypothesis

Given an affine function $f : (\vec{i} \mapsto \dots)$, we have 2 different tilings:

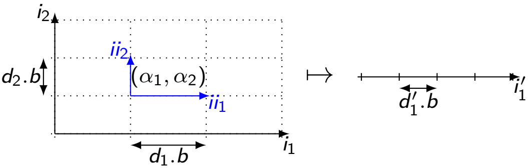
- *Antecedent domain:* $\vec{i} = b.D.\vec{\alpha} + \vec{ii}$ where $\vec{0} \leq \vec{ii} < b.D.\vec{1}$
- *Image domain:* $\vec{i'} = b.D'.\vec{\alpha'} + \vec{ii'}$ where $\vec{0} \leq \vec{ii'} < b.D'.\vec{1}$
- *Parameters:* $\vec{p} = b.\vec{\lambda} + \vec{pp}$ where $\vec{0} \leq \vec{pp} < b.\vec{1}$



Notations and CART hypothesis

Given an affine function $f : (\vec{i} \mapsto \dots)$, we have 2 different tilings:

- *Antecedent domain:* $\vec{i} = b.D.\vec{\alpha} + \vec{ii}$ where $\vec{0} \leq \vec{ii} < b.D.\vec{1}$
- *Image domain:* $\vec{i'} = b.D'.\vec{\alpha'} + \vec{ii'}$ where $\vec{0} \leq \vec{ii'} < b.D'.\vec{1}$
- *Parameters:* $\vec{p} = b.\vec{\lambda} + \vec{pp}$ where $\vec{0} \leq \vec{pp} < b.\vec{1}$



$\Rightarrow$ **Question:** How to obtain ϕ such that $\phi(\vec{\alpha}, \vec{ii}) = (\vec{\alpha'}, \vec{ii'})$?

Deriving the piecewise affine function: polyhedral case

- **Example:** $f : \begin{cases} (i,j) & \mapsto & (2.N + 2.i + 4.j - 1) \\ b \times b & \rightsquigarrow & 2b \end{cases}$

Deriving the piecewise affine function: polyhedral case

- **Example:** $f : \begin{cases} (i,j) & \mapsto & (2.N + 2.i + 4.j - 1) \\ b \times b & \rightsquigarrow & 2b \end{cases}$
- Same computation than for polyhedron (with equalities)
  $\Rightarrow$ We obtain (after derivation):

$$\begin{aligned} \alpha' &= \left\lfloor \frac{2.N_{bl}+2.\alpha+4.\beta}{2} + \frac{2.N_{loc}+2.ii+4.jj-1}{2b} \right\rfloor \\ &= N_{bl} + \alpha + 2\beta + \left\lfloor \frac{2.N_{loc}+2.ii+4.jj-1}{2b} \right\rfloor \end{aligned}$$

Deriving the piecewise affine function: polyhedral case

- **Example:** $f : \begin{cases} (i,j) & \mapsto & (2.N + 2.i + 4.j - 1) \\ b \times b & \rightsquigarrow & 2b \end{cases}$
- Same computation than for polyhedron (with equalities)
  $\Rightarrow$ We obtain (after derivation):

$$\begin{aligned} \alpha' &= \left\lfloor \frac{2.N_{bl}+2.\alpha+4.\beta}{2} + \frac{2.N_{loc}+2.ii+4.jj-1}{2b} \right\rfloor \\ &= N_{bl} + \alpha + 2\beta + \left\lfloor \frac{2.N_{loc}+2.ii+4.jj-1}{2b} \right\rfloor \end{aligned}$$

- $k_1 = \left\lfloor \frac{2.N_{loc}+2.ii+4.jj-1}{2b} \right\rfloor \in [| -1; 3 |]$

Deriving the piecewise affine function: polyhedral case

- **Example:** $f : \begin{cases} (i,j) & \mapsto & (2.N + 2.i + 4.j - 1) \\ b \times b & \rightsquigarrow & 2b \end{cases}$
- Same computation than for polyhedron (with equalities)
  $\Rightarrow$ We obtain (after derivation):

$$\begin{aligned} \alpha' &= \left\lfloor \frac{2.N_{bl}+2.\alpha+4.\beta}{2} + \frac{2.N_{loc}+2.ii+4.jj-1}{2b} \right\rfloor \\ &= N_{bl} + \alpha + 2\beta + \left\lfloor \frac{2.N_{loc}+2.ii+4.jj-1}{2b} \right\rfloor \end{aligned}$$

- $k_1 = \left\lfloor \frac{2.N_{loc}+2.ii+4.jj-1}{2b} \right\rfloor \in [| -1; 3 |]$
- **Result:**

$$(\alpha', ii') = \begin{cases} (N_{bl} + \alpha + 2.\beta - 1, N_{loc} + ii + 2.jj + b) & \text{if} \quad 2.N_{loc} + 2.ii + 4.jj - 1 < 0 \\ (N_{bl} + \alpha + 2.\beta, N_{loc} + ii + 2.jj) & \text{if} \quad 0 \leq 2.N_{loc} + 2.ii + 4.jj - 1 < 2b \\ (N_{bl} + \alpha + 2.\beta + 1, N_{loc} + ii + 2.jj - b) & \text{if} \quad 2b \leq 2.N_{loc} + 2.ii + 4.jj - 1 < 4b \\ (N_{bl} + \alpha + 2.\beta + 2, N_{loc} + ii + 2.jj - 2b) & \text{if} \quad 4b \leq 2.N_{loc} + 2.ii + 4.jj - 1 < 6b \\ (N_{bl} + \alpha + 2.\beta + 3, N_{loc} + ii + 2.jj - 3b) & \text{if} \quad 6b \leq 2.N_{loc} + 2.ii + 4.jj - 1 \end{cases}$$

Deriving the piecewise affine function: general case

- **Example:** $f : \begin{cases} (i) & \mapsto & (i+1) \\ b & \leadsto & 2b \end{cases}$

Deriving the piecewise affine function: general case

- **Example:** $f : \begin{cases} (i) & \mapsto & (i+1) \\ b & \rightsquigarrow & 2b \end{cases}$
- Likewise, we obtain:

$$\alpha' = \left\lfloor \frac{\alpha}{2} + \frac{ii+1}{2b} \right\rfloor$$

Deriving the piecewise affine function: general case

- **Example:** $f : \begin{cases} (i) & \mapsto & (i+1) \\ b & \rightsquigarrow & 2b \end{cases}$
- Likewise, we obtain:

$$\alpha' = \left\lfloor \frac{\alpha}{2} + \frac{ii+1}{2b} \right\rfloor$$

- Let us introduce $\alpha = 2.\mu + \alpha\alpha$, where $0 \leq \alpha\alpha < 2$:

$$\alpha' = \mu + \left\lfloor \frac{\alpha\alpha}{2} + \frac{ii+1}{2b} \right\rfloor$$

- If $\alpha\alpha = 0$, then $k_1 = \left\lfloor \frac{ii+1}{2b} \right\rfloor = 0$.

  If $\alpha\alpha = 1$, then $k_1 = \left\lfloor \frac{b+ii+1}{2b} \right\rfloor = 0$ or 1.

Deriving the piecewise affine function: general case

- **Example:** $f : \begin{cases} (i) & \mapsto & (i+1) \\ b & \rightsquigarrow & 2b \end{cases}$
- Likewise, we obtain:

$$\alpha' = \left\lfloor \frac{\alpha}{2} + \frac{ii+1}{2b} \right\rfloor$$

- Let us introduce $\alpha = 2.\mu + \alpha\alpha$, where $0 \leq \alpha\alpha < 2$:

$$\alpha' = \mu + \left\lfloor \frac{\alpha\alpha}{2} + \frac{ii+1}{2b} \right\rfloor$$

- If $\alpha\alpha = 0$, then $k_1 = \left\lfloor \frac{ii+1}{2b} \right\rfloor = 0$.

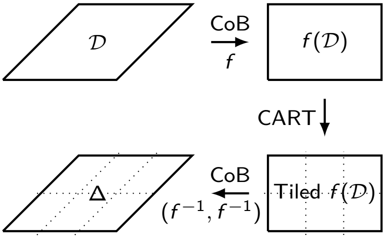
  If $\alpha\alpha = 1$, then $k_1 = \left\lfloor \frac{b+ii+1}{2b} \right\rfloor = 0$ or 1.

- **Result:**

$$(\alpha', ii') = \begin{cases} (\frac{\alpha}{2}, ii+1) & \text{if} \quad \alpha \bmod 2 = 0 \\ (\frac{\alpha-1}{2}, b+ii+1) & \text{if} \quad \alpha \bmod 2 = 1 \ \wedge \ b+ii+1 < 2b \\ (\frac{\alpha-1}{2}+1, ii+1-b) & \text{if} \quad \alpha \bmod 2 = 1 \ \wedge \ 2b \leq b+ii+1 \end{cases}$$

Extensions of CART

- **CART along non-canonic axis:**



- **Several tile size parameters:**
  - Works if the tile size parameters do not interfere.
    - $\rightarrow$ Ex: matrix multiply with 3 tile size parameters.
  - Else, $\left\lfloor \frac{b}{b'} \right\rfloor$: not manageable with the same kind of technique.

Conclusion and future works

- Code still polyhedral after the CART transformation
  - ⇒ Allow polyhedral optimization/analysis after parametric tiling
  - ⇒ Pluggable in the compilation flow for free
- Library standalone implementation available in Java/C++
  - → http://compsys-tools.ens-lyon.fr/
- Implementation of the full CART transformation in the *AlphaZ framework* in progress
- Use CART as the first step of semantic tiling transformation
  - Extract a program piece which uses a specific tile of data
  - → Recognize this piece as an higher-order operator

Thank you for listening

Do you have any questions?