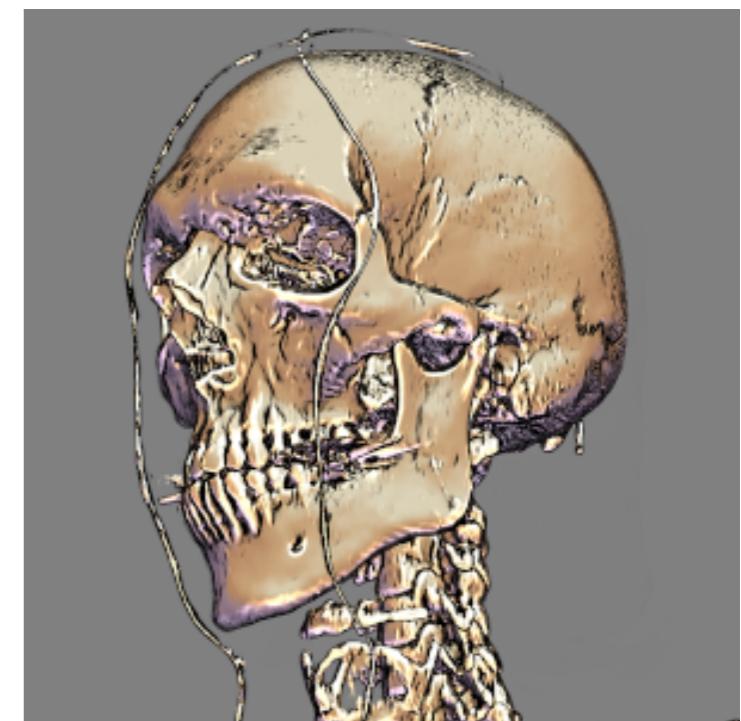


Rewriting with an index-based intermediate representation

Charisee Chiw and John Reppy

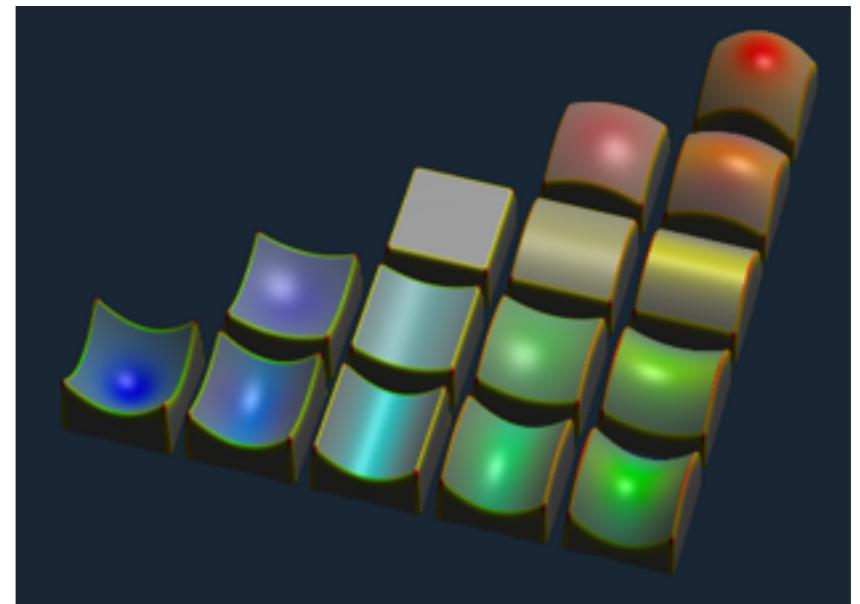
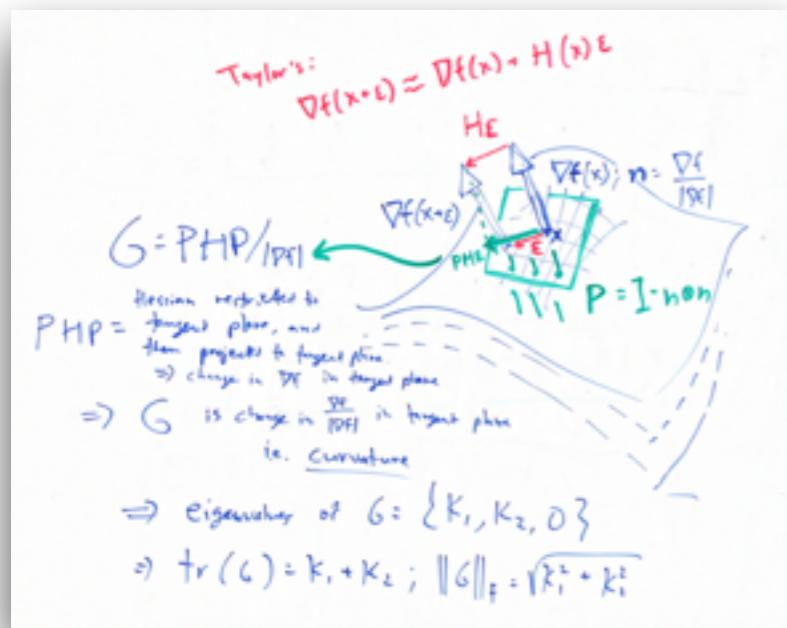


Roadmap

- Domain Space
- EIN operators as a compact representation of loops over tensor indices (used for tensor math)
- Index-based optimizations
- Potential for loop optimization

Domain

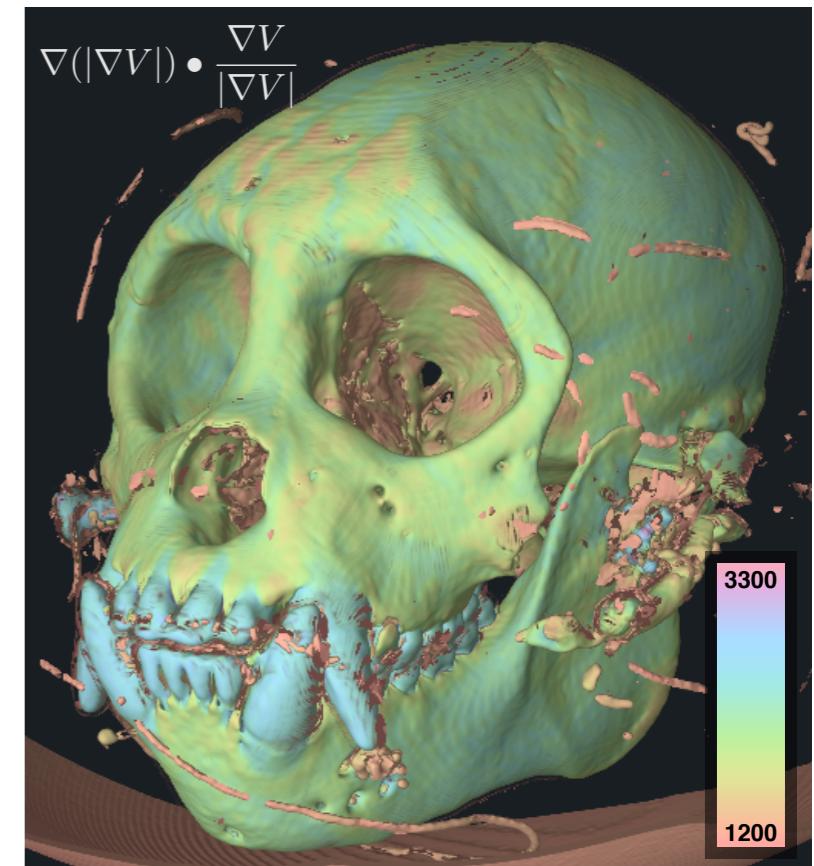
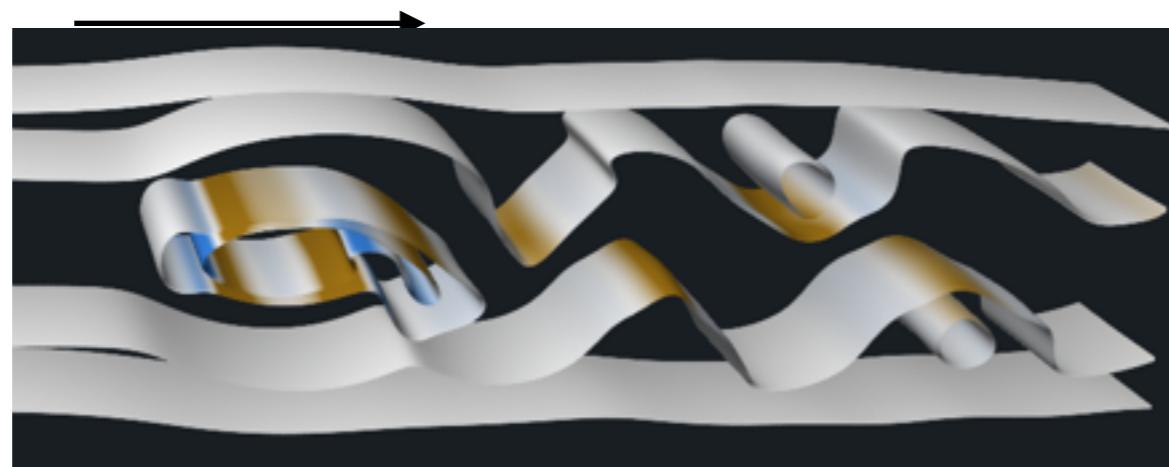
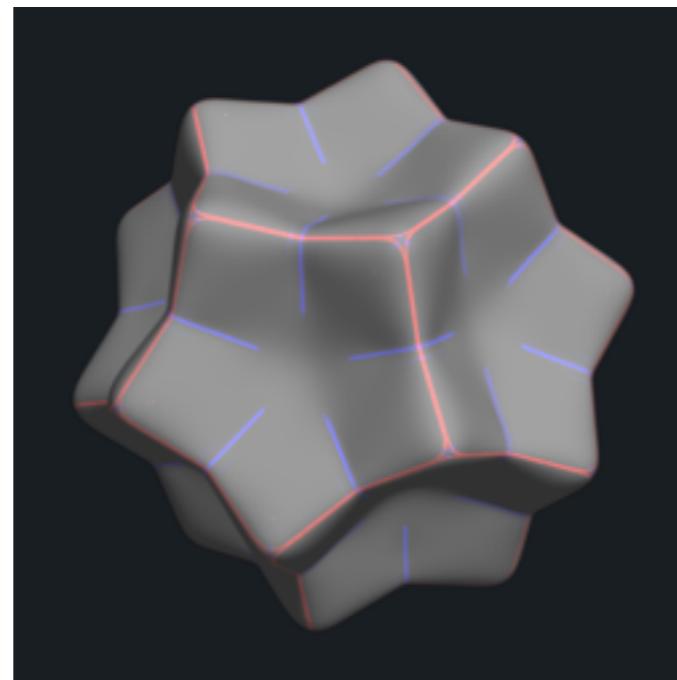
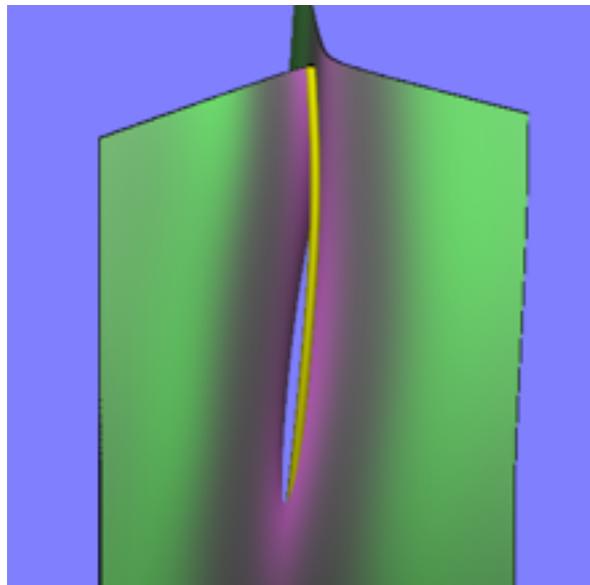
- Diderot; DSL for scivis and image analysis [PLDI'12]
- Scientists use algorithms to explore image space and understand data
- Expressive language for tensor and field computations



Domain

Sophisticated Visualization algorithms

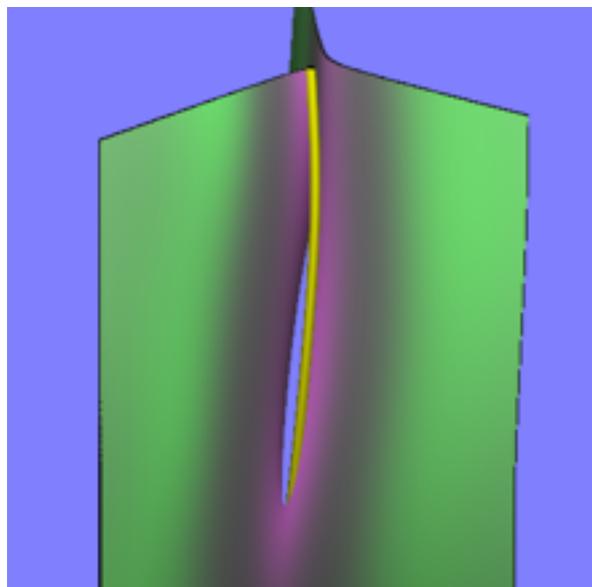
Algorithm in this space require a lot of tensor math



Domain

Implementing Sophisticated Visualization algorithms

Visualization



computation

$$E = \nabla \otimes \nabla F - \text{identity}[3] * \text{trace}(\nabla \otimes \nabla F)/3;$$

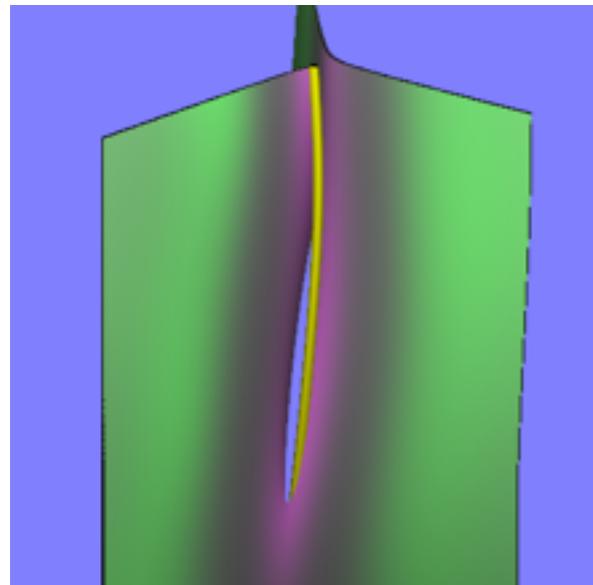
$$D = 3 \sqrt{6} * \det(E/|E|);$$

$$T = \nabla \otimes \nabla D$$

Domain

Implementing Sophisticated Visualization algorithms

Visualization



User does derivations

first-order higher-order

$$\begin{aligned} q &= \text{delV}(x) \\ \text{deltr} &= q[0,0,:] + q[1,1,:] + q[2,2,:] \\ E &= V(x) - \text{trace}(\text{identity}[3]) / 3 \\ \text{val} &= 3\sqrt{6} \cdot \det(E / \sqrt{E:E}) \\ K &= (\text{identity}[3] \otimes \text{deltr}) / 3 \\ \text{delE} &= \text{delV}(x) - K \quad E = V(x) - \text{trace}(V) \cdot \text{identity}[3] \\ N1 &= \text{delE} \cdot \sqrt{E:E} \quad F = 3\sqrt{6} \cdot \det(E / |E|) \\ \text{inner} &= 2 \cdot E \cdot \text{delE} \quad a = \alpha(F(x), |\nabla F(x)|) \\ B &= (\text{inner} / (E:E)) / 2 \\ N2 &= E \otimes B \\ dQ &= (N1 - N2) / E:E \\ Q &= E / \sqrt{E:E} \\ \text{part1} &= dQ[0,0] \cdot (Q[1,1] \cdot Q[2,2] - Q[1,2] \cdot Q[2,1]) \\ &\quad + Q[0,0] \cdot (Q[2,2] \cdot dQ[1,1,:] + Q[1,1] \cdot dQ[2,2,:]) \\ &\quad - dQ[1,2] \cdot Q[2,1] + Q[1,2] \cdot dQ[2,1,:]) \\ \text{part2} &= dQ[0,1,:] \cdot (Q[1,0] \cdot Q[2,2] - Q[1,2] \cdot Q[2,0]) \\ &\quad + Q[0,1] \cdot (Q[2,2] \cdot dQ[1,0,:] + Q[1,0] \cdot dQ[2,2,:]) \\ &\quad - dQ[1,2] \cdot Q[2,0] + Q[1,2] \cdot dQ[2,0,:]) \\ \text{part3} &= dQ[0,2,:] \cdot (Q[0,0] \cdot Q[2,1] - Q[0,1] \cdot Q[2,0]) \\ &\quad + Q[0,2] \cdot (Q[2,1] \cdot Q[1,0] + Q[1,0] \cdot dQ[2,1,:]) \\ &\quad - dQ[1,1,:] \cdot Q[2,1] + Q[1,1] \cdot dQ[2,1,:]) \\ \text{delM} &= \text{part1} - \text{part2} + \text{part3} \\ \text{grad} &= 3\sqrt{6} \cdot \text{delM} \\ a &= \alpha(\text{val}, \text{lgrad}) \end{aligned}$$

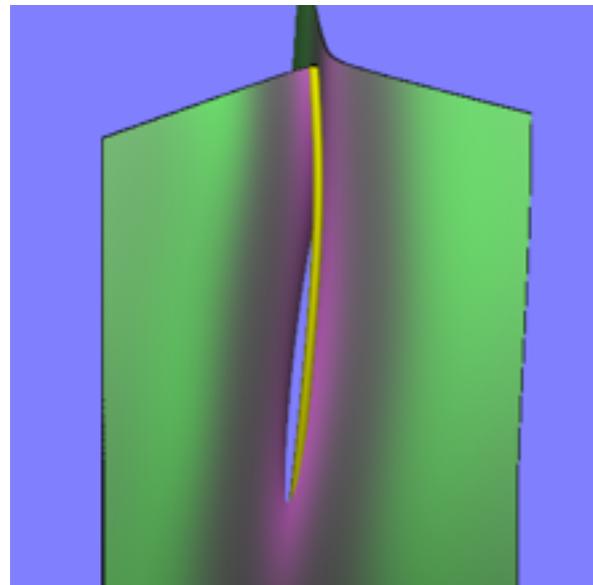
computation

$$\begin{aligned} E &= \nabla \otimes \nabla F - \text{identity}[3] * \text{trace}(\nabla \otimes \nabla F) / 3; \\ D &= 3 \sqrt{6} * \det(E / |E|); \\ T &= \nabla \otimes \nabla D \end{aligned}$$

Domain

Implementing Sophisticated Visualization algorithms

Visualization



computation

$$E = \nabla \otimes \nabla F - \text{identity}[3] * \text{trace}(\nabla \otimes \nabla F) / 3;$$
$$D = 3 \sqrt{6} * \det(E / |E|);$$
$$T = \nabla \otimes \nabla D$$

User does derivations

first-order higher-order

$$q = \text{delV}(x)$$
$$\text{deltr} = q[0,0,:] + q[1,1,:] + q[2,2,:]$$
$$E = V(x) - \text{trace} * \text{identity}[3] / 3$$
$$\text{val} = 3 \cdot \sqrt{6} \cdot \det(E / \sqrt{E:E})$$
$$K = (\text{identity}[3] \otimes \text{deltr}) / 3$$
$$\text{delE} = \text{delV}(x) - K$$
$$E = V(x) - \text{trace}(V) \cdot \text{identity}[3]$$
$$F = 3 \cdot \sqrt{6} \cdot \det(E / |E|)$$
$$N1 = \text{delE} \cdot \sqrt{E:E}$$
$$\alpha = \text{alpha}(F(x), |\nabla F(x)|)$$
$$\text{inner} = 2 \cdot E \cdot \text{delE}$$
$$B = (\text{inner} / (E:E)) / 2$$
$$N2 = E \otimes B$$
$$dQ = (N1 - N2) / E:E$$
$$Q = E / \sqrt{E:E}$$
$$\text{part1} = dQ[0,0] * (Q[1,1]Q[2,2] - Q[1,2]Q[2,1])$$
$$+ Q[0,0] * (Q[2,2] * dQ[1,1,:] + Q[1,1] * dQ[2,2,:])$$
$$- dQ[1,2] * Q[2,1] + Q[1,2] * dQ[2,1,:]$$
$$\text{part2} = dQ[0,1,:] * (Q[1,0] * Q[2,2] - Q[1,2] * Q[2,0])$$
$$+ Q[0,1] * (Q[2,2] * dQ[1,0,:] + Q[1,0] * dQ[2,2,:])$$
$$- dQ[1,2] * Q[2,0] + Q[1,2] * dQ[2,0,:]$$
$$\text{part3} = dQ[0,2,:] * (Q[0,0] * Q[2,1] - Q[0,1] * Q[2,0])$$
$$Q[0,2] * (Q[2,1] * Q[1,0] + Q[1,0] * dQ[2,1,:])$$
$$- dQ[1,1,:] * Q[2,1] + Q[1,1] * dQ[2,1,:]$$
$$\text{delM} = \text{part1} - \text{part2} + \text{part3}$$
$$\text{grad} = 3 \cdot \sqrt{6} \cdot \text{delM}$$
$$\alpha = \text{alpha}(\text{val}, |\text{grad}|)$$

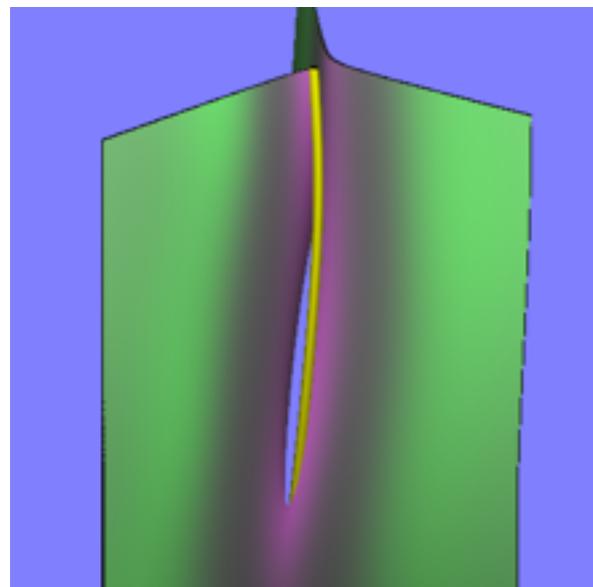
Diderot Program '12



Domain

Implementing Sophisticated Visualization algorithms

Visualization



New and improved Diderot program



computation

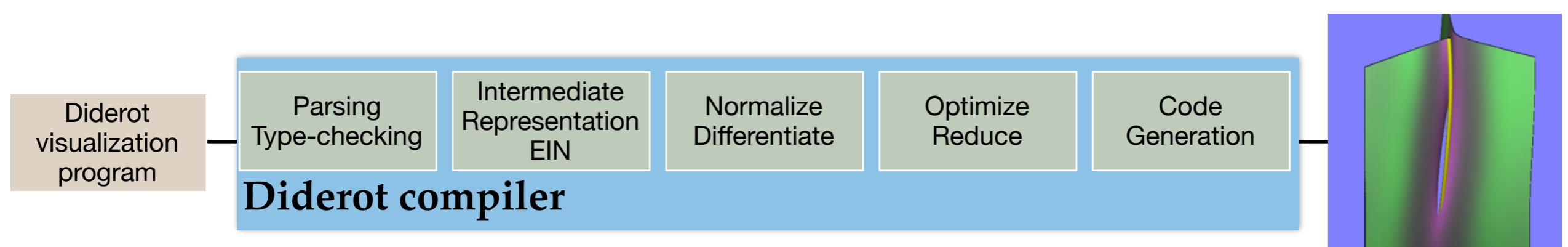
$$E = \nabla \otimes \nabla F - \text{identity}[3] * \text{trace}(\nabla \otimes \nabla F) / 3;$$

$$D = 3 \sqrt{6} * \det(E / |E|);$$

$$T = \nabla \otimes \nabla D$$

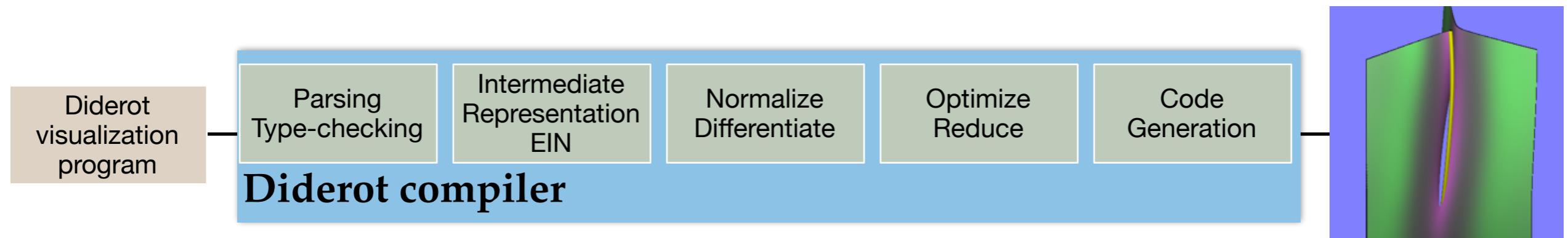
fourth derivatives

From Tensor calculus to C



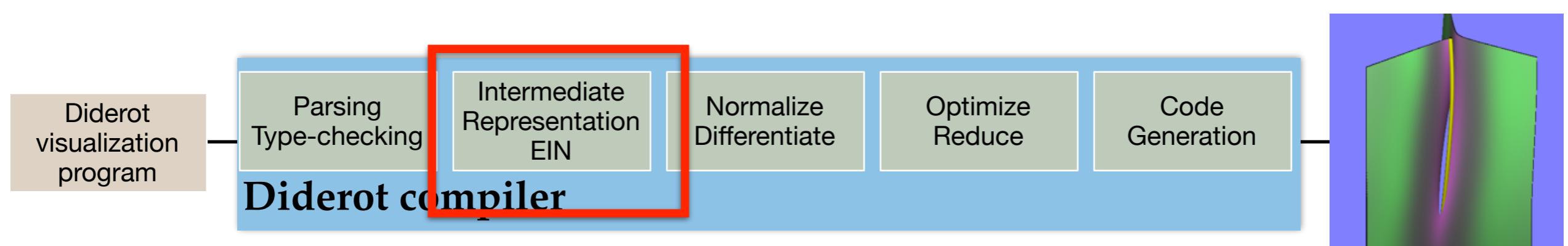
From Tensor calculus to C

- Diderot offers ability to write computations at a high level



From Tensor calculus to C

- Diderot offers ability to write computations at a high level
- Enabled by design and implementation of EIN IR [Masters '14]



Rewriting with EIN

- Hybrid design that embeds expression trees into a normalized SSA representation

$X \ Y \ Z$

Rewriting with EIN

- Hybrid design that embeds expression trees into a normalized SSA representation

$$X_\alpha Y_\beta Z_\beta$$

- Indices on tensor and field argument illustrate sampling

Rewriting with EIN

- Hybrid design that embeds expression trees into a normalized SSA representation

$$\sum_{\beta} X_{\alpha} Y_{\beta} Z_{\beta}$$

- Indices on tensor and field argument illustrate sampling
- **Summation** operator represents a loop nest that will be unrolled later

Rewriting with EIN

- Hybrid design that embeds expression trees into a normalized SSA representation

$$\sum_{\beta} X_{\alpha} Y_{\beta} Z_{\beta} \rightarrow X_{\alpha} \sum_{\beta} Y_{\beta} Z_{\beta}$$

- Indices on tensor and field argument illustrate sampling
- **Summation** operator represents a loop nest that will be unrolled later
- Moving terms outside can avoid subsequent code duplication

Intermediate Representation: EIN

The matrix M is represented in EIN with two variable indices

$$\text{Matrix } M \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} = M_{ij} \quad \text{EIN Term}$$

Intermediate Representation: EIN

The matrix M is represented in EIN with two variable indices

Matrix M
$$\begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} = M_{ij}$$
 EIN Term

Switch row and column
$$\begin{bmatrix} M_{11} & M_{21} \\ M_{12} & M_{22} \end{bmatrix} = M_{ji}$$
 Switch the order

Intermediate Representation: EIN

The matrix M is represented in EIN with two variable indices

Matrix M
$$\begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} = M_{ij}$$
 EIN Term

Switch row and column
$$\begin{bmatrix} M_{11} & M_{21} \\ M_{12} & M_{22} \end{bmatrix} = M_{ji}$$
 Switch the order

Intermediate Representation: EIN

The matrix M is represented in EIN with two variable indices

Matrix M
$$\begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} = M_{ij}$$
 EIN Term

Switch row and column
$$\begin{bmatrix} M_{11} & M_{21} \\ M_{12} & M_{22} \end{bmatrix} = M_{ji}$$
 Switch the order

Select a column
$$\begin{bmatrix} M_{11} \\ M_{21} \end{bmatrix} = M_{i1}$$
 Mix of different indices

Single entry
$$M_{12} = M_{12}$$
 Set indices to constants

Add diagonal components
$$\text{trace}(M) = \sum_i M_{ii}$$
 Summation

Intermediate Representation: EIN

The matrix M is represented in EIN with two variable indices

$$\text{Matrix } M \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} = M_{ij} \quad \text{EIN Term}$$

Intermediate Representation: EIN

The matrix M is represented in EIN with two variable indices

$$\text{Matrix } M \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} = M_{ij} \quad \text{EIN Term}$$

Matrix Inverse

$$\text{inv}(M) = \frac{\sum_k M_{kk} \delta_{ij} - M_{ij}}{M_{00}M_{11} - M_{01}M_{10}}$$

Matrix Determinant

$$\det(M) = \sum_{ijk} M_{0i} M_{1j} M_{2k} \mathcal{E}_{ijk}$$

Intermediate Representation: EIN

- Operation between terms

$$\sum_i A_i B_i$$

$$A_i B_i$$

$$A_i B_j$$

Intermediate Representation: EIN

- Operation between terms

real	vector	matrix
$\sum_i A_i B_i$ dot product	$A_i B_i$ component-wise multiplication	$A_i B_j$ outer product

Intermediate Representation: EIN

- Operation between terms

$$\sum_i A_i B_i$$

dot product

real

vector

matrix

$$A_i B_i$$

component-wise multiplication

$$A_i B_j$$

outer product

- Can apply mix of different operators to each other:
composition, curl, concatenation, Hessian, Transpose, Double dot product

$$\nabla(F \circ G)$$

$$Cons[T[:0], T[1:]]$$

$$\nabla \times$$

$$\nabla \otimes \nabla((A : B) \times C)$$

Intermediate Representation: EIN

- Operation between terms

real	vector	matrix
$\sum_i A_i B_i$ dot product	$A_i B_i$ component-wise multiplication	$A_i B_j$ outer product

- Can apply mix of different operators to each other:
composition, curl, concatenation, Hessian, Transpose, Double dot product

$$\nabla(F \circ$$

EIN supports a rich language for tensor
and field computations

$$Cons[T[:0], T[1:]]$$

$$\nabla \times$$

$$\nabla \otimes \nabla((A : B) \times C)$$

Intermediate Representation: EIN

Consider the computation

$$t = \text{trace}(a \otimes b)$$

Intermediate Representation: EIN

Consider the computation

$$t = \text{trace}(a \otimes b)$$

Diderot expression is represented in compiler as

$$M = (\lambda(A, B) \langle A_i B_j \rangle_{ij})(a, b)$$

$$t = (\lambda T \langle \sum_i T_{ii} \rangle)(M)$$

Intermediate Representation: EIN

Consider the computation

$$t = \text{trace}(a \otimes b)$$

Diderot expression is represented in compiler as

$$\begin{aligned} M &= (\lambda(A, B) \langle A_i B_j \rangle_{ij})(a, b) \\ t &= (\lambda T \langle \sum_i T_{ii} \rangle)(M) \end{aligned}$$

substitution of the definition of M for X yields

$$t = (\lambda(A, B) \langle \sum_i A_i B_i \rangle)(a, b)$$

Intermediate Representation: EIN

Consider the computation

$$t = \text{trace}(a \otimes b)$$

Diderot expression is represented in compiler as

$$\begin{aligned} M &= (\lambda(A, B) \langle A_i B_j \rangle_{ij})(a, b) \\ t &= (\lambda T \langle \sum_i T_{ii} \rangle)(M) \end{aligned}$$

substitution of the definition of M for X yields

$$t = (\lambda(A, B) \langle \sum_i A_i B_i \rangle)(a, b)$$

Replaces the ad hoc rewrite rule:

$$\text{Trace}(\text{Outer}(a, b)) \Rightarrow \text{Dot}(a, b)$$

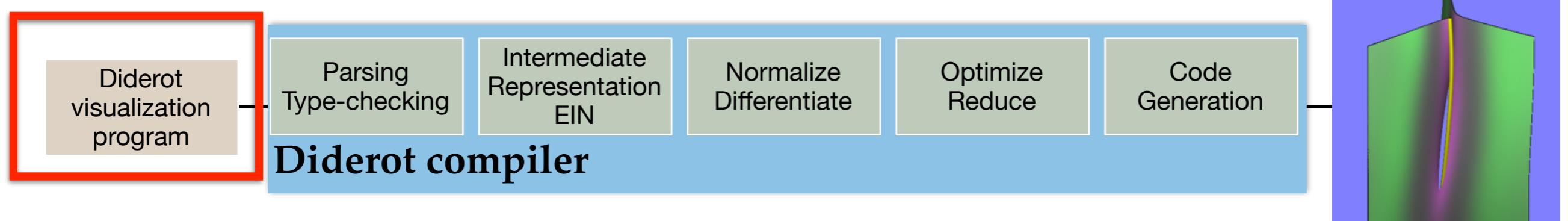
From Tensor calculus to C

Core Computation in visualization program

$$E = \nabla \otimes \nabla F - \text{identity}[3] * \text{trace}(\nabla \otimes \nabla F) / 3;$$

$$D = 3 \sqrt{6} * \det(E / |E|);$$

$$T = \nabla \otimes \nabla D$$



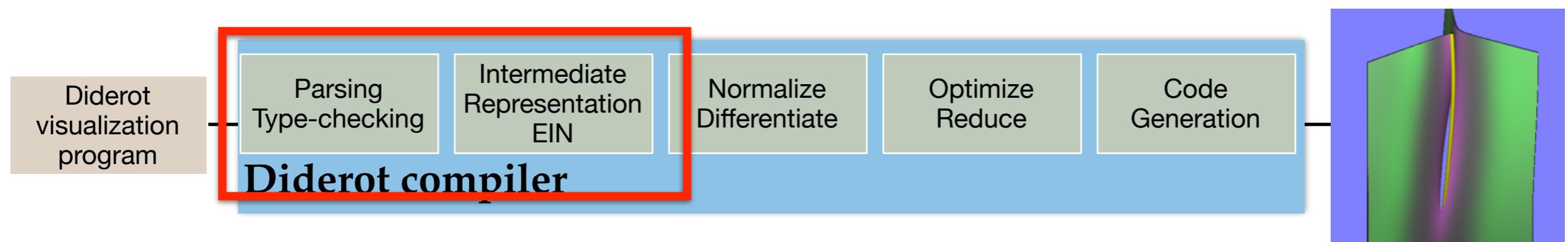
From Tensor calculus to C

Core Computation in visualization program

$$E = \nabla \otimes \nabla F - \text{identity}[3] * \text{trace}(\nabla \otimes \nabla F) / 3;$$

$$D = 3 \sqrt{6} * \det(E / |E|);$$

$$T = \nabla \otimes \nabla D$$



From Tensor calculus to C

Core Computation in visualization program

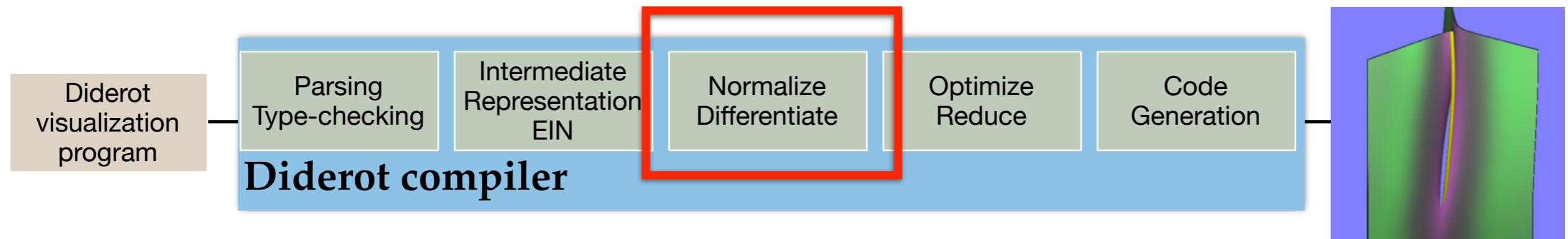
$$E = \nabla \otimes \nabla F - \text{identity}[3] * \text{trace}(\nabla \otimes \nabla F) / 3;$$

$$D = 3 \sqrt{6} * \det(E / |E|);$$

$$T = \nabla \otimes \nabla D$$

Necessary Rewriting

$$\nabla |\nabla F|(x) \longrightarrow \frac{1}{2} \frac{(\nabla \otimes \nabla F(p) \bullet \nabla F(p)) + (\nabla F(p) \bullet \nabla \otimes \nabla F(p))}{\sqrt{\nabla F(p) \bullet \nabla F(p)}}$$



From Tensor calculus to C

Core Computation in visualization program

$$E = \nabla \otimes \nabla F - \text{identity}[3] * \text{trace}(\nabla \otimes \nabla F) / 3;$$

$$D = 3 \sqrt{6} * \det(E / |E|);$$

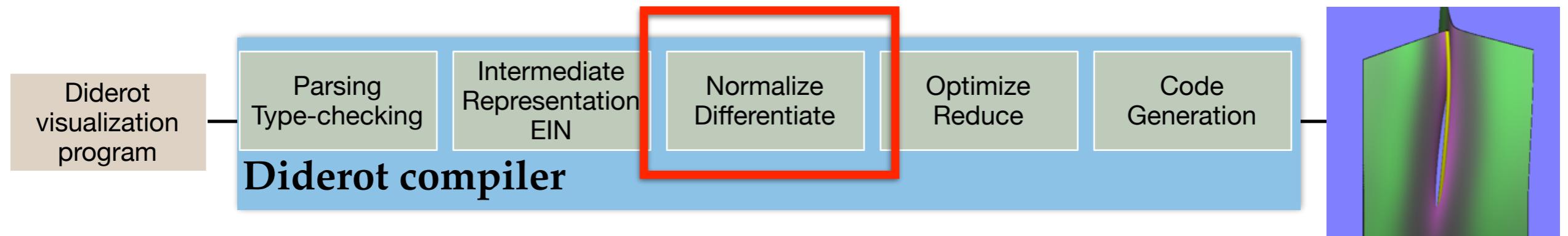
$$T = \nabla \otimes \nabla D$$

Necessary Rewriting

$$\nabla |\nabla F|(x) \longrightarrow \frac{1}{2} \frac{(\nabla \otimes \nabla F(p) \bullet \nabla F(p)) + (\nabla F(p) \bullet \nabla \otimes \nabla F(p))}{\sqrt{\nabla F(p) \bullet \nabla F(p)}}$$

The Issue

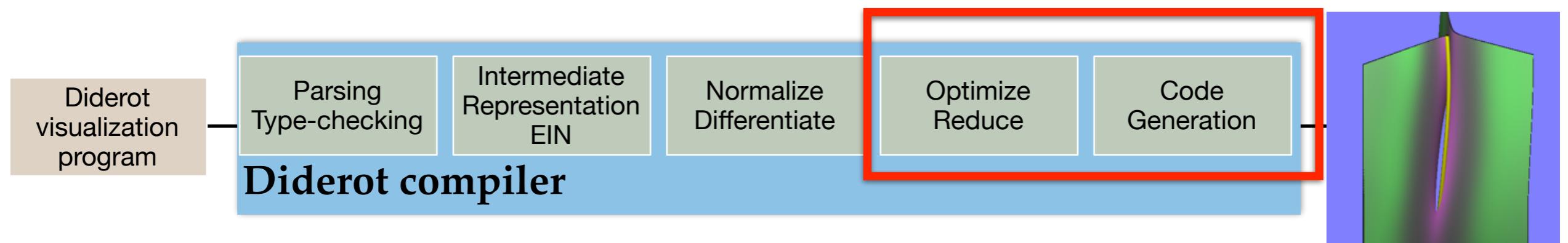
- Rich Language + necessary rewriting = bulky obtuse IR
- It becomes difficult to compile and generate efficient code



From Tensor calculus to C

Index Based Rewriting

- Shift invariant terms
- Split EIN computations into SSA representation based on EIN structure
- Examine Structure and vectorize



1. Shift invariant terms

User writes high-level computation

$$\text{out} = (s^*(v \otimes x)) \bullet u$$

1. Shift invariant terms

User writes high-level computation

$$\text{out} = (\mathbf{s}^* (\mathbf{v} \otimes \mathbf{x})) \bullet \mathbf{u}$$

Internally represented with EIN operator

$$\text{out} = \left\langle \sum_j (s v_i x_j u_j) \right\rangle_i$$

1. Shift invariant terms

User writes high-level computation

$$\text{out} = (\mathbf{s}^*(\mathbf{v} \otimes \mathbf{x})) \bullet \mathbf{u}$$

Internally represented with EIN operator

$$\text{out} = \left\langle \sum_j (sv_i x_j u_j) \right\rangle_i \xrightarrow{\text{Shift}} \left\langle sv_i \sum_j (x_j u_j) \right\rangle_i$$

1. Shift invariant terms

User writes high-level computation

$$\text{out} = (\mathbf{s}^*(\mathbf{v} \otimes \mathbf{x})) \bullet \mathbf{u}$$

Internally represented with EIN operator

$$\text{out} = \left\langle \sum_j (sv_i x_j u_j) \right\rangle_i \xrightarrow{\text{Shift}} \left\langle sv_i \sum_j (x_j u_j) \right\rangle_i$$

Operators are split apart based on EIN structure

$$\begin{array}{c} \text{Split} \\ \Rightarrow \end{array} \quad \begin{array}{l} \text{out} = \langle sv_i d \rangle_i \\ d = \left\langle \sum_j x_j u_j \right\rangle \end{array}$$

1. Shift invariant terms

User writes high-level computation

$$\text{out} = (\mathbf{s}^*(\mathbf{v} \otimes \mathbf{x})) \bullet \mathbf{u}$$

Internally represented with EIN operator

$$\text{out} = \left\langle \sum_j (sv_i x_j u_j) \right\rangle_i \xrightarrow{\text{Shift}} \left\langle sv_i \sum_j (x_j u_j) \right\rangle_i$$

Operators are split apart based on EIN structure

$$\begin{array}{c} \text{Split} \\ \Rightarrow \end{array} \quad \begin{array}{l} \text{out} = \langle sv_i d \rangle_i \\ d = \underbrace{\left\langle \sum_j x_j u_j \right\rangle}_{\text{vector operation}} \end{array}$$

1. Shift invariant terms

User writes high-level computation

$$\text{out} = (\mathbf{s}^*(\mathbf{v} \otimes \mathbf{x})) \bullet \mathbf{u} \rightarrow \mathbf{s}^*(\mathbf{v}^*(\mathbf{x} \bullet \mathbf{u}))$$

Replace a direct-style rewrite

Internally represented with EIN operator

$$\text{out} = \left\langle \sum_j (sv_i x_j u_j) \right\rangle_i \xrightarrow{\text{Shift}} \left\langle sv_i \sum_j (x_j u_j) \right\rangle_i$$

Operators are split apart based on EIN structure

$$\begin{array}{c} \xrightarrow{\text{Split}} \\ \Rightarrow \end{array} \quad \begin{array}{l} \text{out} = \langle sv_i d \rangle_i \\ d = \underbrace{\langle \sum_j x_j u_j \rangle}_{\text{vector operation}} \end{array}$$

2. Separate a large loop

Rewrites allow us to eliminate summations

(i.e., effectively reduce the loop-nesting depth)

$$(a \times b) \bullet (c \times d)$$

User writes high-level computation

2. Separate a large loop

Rewrites allow us to eliminate summations

(i.e., effectively reduce the loop-nesting depth)

$$(a \times b) \bullet (c \times d)$$

User writes high-level computation

$$\Rightarrow \sum_i (\sum_{jk} \epsilon_{ijk} a_j b_k) (\sum_{lm} \epsilon_{ilm} c_l d_m)$$

EIN operator

2. Separate a large loop

Rewrites allow us to eliminate summations

(i.e., effectively reduce the loop-nesting depth)

$$(a \times b) \bullet (c \times d)$$

User writes high-level computation

$$\Rightarrow \sum_i (\sum_{jk} \boxed{\mathcal{E}_{ijk}} a_j b_k) (\sum_{lm} \boxed{\mathcal{E}_{ilm}} c_l d_m)$$

EIN operator

2. Separate a large loop

Rewrites allow us to eliminate summations

(i.e., effectively reduce the loop-nesting depth)

$$(a \times b) \bullet (c \times d)$$

User writes high-level computation

$$\Rightarrow \cancel{\sum_i} (\sum_{jk} \boxed{\epsilon_{ijk}} a_j b_k) (\sum_{lm} \boxed{\epsilon_{ilm}} c_l d_m)$$

EIN operator

$$\Rightarrow (\sum_{jklm} \delta_{jl} \delta_{km} a_j b_k c_l d_m) - (\sum_{jklm} \delta_{jm} \delta_{kl} a_j b_k c_l d_m)$$

Epsilons with shared index

2. Separate a large loop

Rewrites allow us to eliminate summations

(i.e., effectively reduce the loop-nesting depth)

$$(a \times b) \bullet (c \times d)$$

User writes high-level computation

$$\Rightarrow \cancel{\sum_i} (\sum_{jk} \boxed{\epsilon_{ijk}} a_j b_k) (\sum_{lm} \boxed{\epsilon_{ilm}} c_l d_m) \quad \text{EIN operator}$$

$$\Rightarrow (\sum_{jklm} \boxed{\delta_{jl}\delta_{km}} a_j b_k c_l d_m) - \sum_{jklm} \boxed{\delta_{jm}\delta_{kl}} a_j b_k c_l d_m \quad \text{Epsilons with shared index}$$

2. Separate a large loop

Rewrites allow us to eliminate summations

(i.e., effectively reduce the loop-nesting depth)

$$(a \times b) \bullet (c \times d)$$

User writes high-level computation

$$\Rightarrow \sum_i (\sum_{jk} \epsilon_{ijk} a_j b_k) (\sum_{lm} \epsilon_{ilm} c_l d_m)$$

EIN operator

$$\Rightarrow (\sum_{jklm} \delta_{jl} \delta_{km} a_j b_k c_l d_m) - \sum_{jklm} \delta_{jm} \delta_{kl} a_j b_k c_l d_m)$$

Epsilons with shared index

$$\Rightarrow (\sum_{jk} a_j b_k c_j d_k) - \sum_{jk} (a_j b_k c_k d_j)$$

Delta Reduction

2. Separate a large loop

Rewrites allow us to eliminate summations

(i.e., effectively reduce the loop-nesting depth)

$$(a \times b) \bullet (c \times d)$$

User writes high-level computation

$$\Rightarrow \sum_i (\sum_{jk} \epsilon_{ijk} a_j b_k) (\sum_{lm} \epsilon_{ilm} c_l d_m)$$

EIN operator

$$\Rightarrow (\sum_{jklm} \delta_{jl} \delta_{km} a_j b_k c_l d_m) - (\sum_{jklm} \delta_{jm} \delta_{kl} a_j b_k c_l d_m)$$

Epsilons with shared index

$$\Rightarrow (\sum_{jk} a_j b_k c_j d_k) - (\sum_{jk} a_j b_k c_k d_j)$$

Delta Reduction

Shift

$$\Rightarrow (\sum_j a_j c_j) (\sum_k b_k d_k) - (\sum_j a_j d_j) (\sum_k b_k c_k)$$

Summation binding

2. Separate a large loop

Rewrites allow us to eliminate summations

(i.e., effectively reduce the loop-nesting depth)

$$(a \times b) \bullet (c \times d)$$

User writes high-level computation

$$\Rightarrow \sum_i (\sum_{jk} \epsilon_{ijk} a_j b_k) (\sum_{lm} \epsilon_{ilm} c_l d_m)$$

EIN operator

$$\Rightarrow (\sum_{jklm} \delta_{jl} \delta_{km} a_j b_k c_l d_m) - (\sum_{jklm} \delta_{jm} \delta_{kl} a_j b_k c_l d_m)$$

Epsilons with shared index

$$\Rightarrow (\sum_{jk} a_j b_k c_j d_k) - (\sum_{jk} a_j b_k c_k d_j)$$

Delta Reduction

Shift

$$\Rightarrow (\sum_j a_j c_j) (\sum_k b_k d_k) - (\sum_j a_j d_j) (\sum_k b_k c_k)$$

Summation binding

vector operation

2. Separate a large loop

Rewrites allow us to eliminate summations

(i.e., effectively reduce the loop-nesting depth)

$(a \times b) \bullet (c \times d) \Rightarrow (a \bullet c)(b \bullet d) - (a \bullet d)(b \bullet c)$ User writes high-level computation

Replace a direct-style rewrite

$$\Rightarrow \cancel{\sum_i} (\sum_{jk} \epsilon_{ijk} a_j b_k) (\sum_{lm} \epsilon_{ilm} c_l d_m) \quad \text{EIN operator}$$

$$\Rightarrow (\sum_{\cancel{jklm}} \delta_{jl} \delta_{km} a_j b_k c_l d_m) - \sum_{\cancel{jklm}} (\delta_{jm} \delta_{kl} a_j b_k c_l d_m) \quad \text{Epsilons with shared index}$$

$$\Rightarrow (\sum_{jk} a_j b_k c_j d_k) - \sum_{jk} (a_j b_k c_k d_j) \quad \text{Delta Reduction}$$

Shift

$$\Rightarrow (\sum_j a_j c_j) (\sum_k b_k d_k) - (\sum_j a_j d_j) (\sum_k b_k c_k) \quad \text{Summation binding}$$

vector operation

3. Outer loop decision

User writes high-level computation $(c \bullet d^*a) \bullet b$

3. Outer loop decision

User writes high-level computation $(c \bullet d^* a) \bullet b$

Internally represented with EIN operator

$$e = \sum_{ij} AB_i C_j D_{ji}$$

3. Outer loop decision

User writes high-level computation $(c \bullet d^* a) \bullet b$

Internally represented with EIN operator

$$e = \sum_{ij} AB_i C_j D_{ji}$$

Two possible rewrites. Which expression is best?

$$e \xrightarrow{R2} A \sum_j C_j \sum_i D_{ji} B_i$$

$$e \xrightarrow{R1} A \sum_i (B_i \sum_j C_j D_{ji})$$

3. Outer loop decision

User writes high-level computation $(c \bullet d^* a) \bullet b$

Internally represented with EIN operator

$$e = \sum_{ij} AB_i C_j D_{ji}$$

Two possible rewrites. Which expression is best?

$$e \xrightarrow{R2} A \sum_j C_j \sum_i \underline{D_{ji}} B_i$$

Row Projection

$$e \xrightarrow{R1} A \sum_i (B_i \sum_j \underline{C_j} D_{ji})$$

Column Projection

3. Outer loop decision

User writes high-level computation $(c \bullet d^*a) \bullet b$

Internally represented with EIN operator

$$e = \sum_{ij} AB_i C_j D_{ji}$$

Two possible rewrites. Which expression is best?

$$e \xrightarrow{R2} A \sum_j C_j \sum_i \underline{D_{ji} B_i}$$

Row Projection

$$e \xrightarrow{R1} A \sum_i (B_i \sum_j \underline{C_j D_{ji}})$$

Column Projection

Split $t_0 = \lambda(f, d) \langle \sum_i D_{ji} B_i \rangle_j (f, d)$

$$\text{out} = \lambda(c, e, t) \langle A \sum_j C_j t_j \rangle (e, t_0)$$

3. Outer loop decision

User writes high-level computation $(c \bullet d^*a) \bullet b$

Internally represented with EIN operator

$$e = \sum_{ij} AB_i C_j D_{ji}$$

Two possible rewrites. Which expression is best?

$$e \xrightarrow{R2} A \sum_j C_j \sum_i \underline{D_{ji}} B_i$$

Row Projection

$$e \xrightarrow{R1} A \sum_i (B_i \sum_j \underline{C_j} D_{ji})$$

Column Projection

Split $t_0 = \lambda(f, d) \left\langle \sum_i D_{ji} B_i \right\rangle_j (f, d)$

$$\text{out} = \lambda(c, e, t) \left\langle A \sum_j \underline{C_j} t_j \right\rangle (e, t_0)$$

vector operation

The big question

- Can we expand on this approach?
- Is splitting based on the EIN structure the best thing?
- Could we take a step back and look at bigger mathematical structure?
- Performance?

Conclusion

- EIN operators as a compact representation of loops over tensor indices (used for tensor math)
- Invariant terms can be moved in/out of loops with simple pattern matching, rewriting, and analysis
- In the future we want to examine the entire computation earlier in the computation

Thank you

Github Page

<https://github.com/Diderot-Language/examples>

Diderot

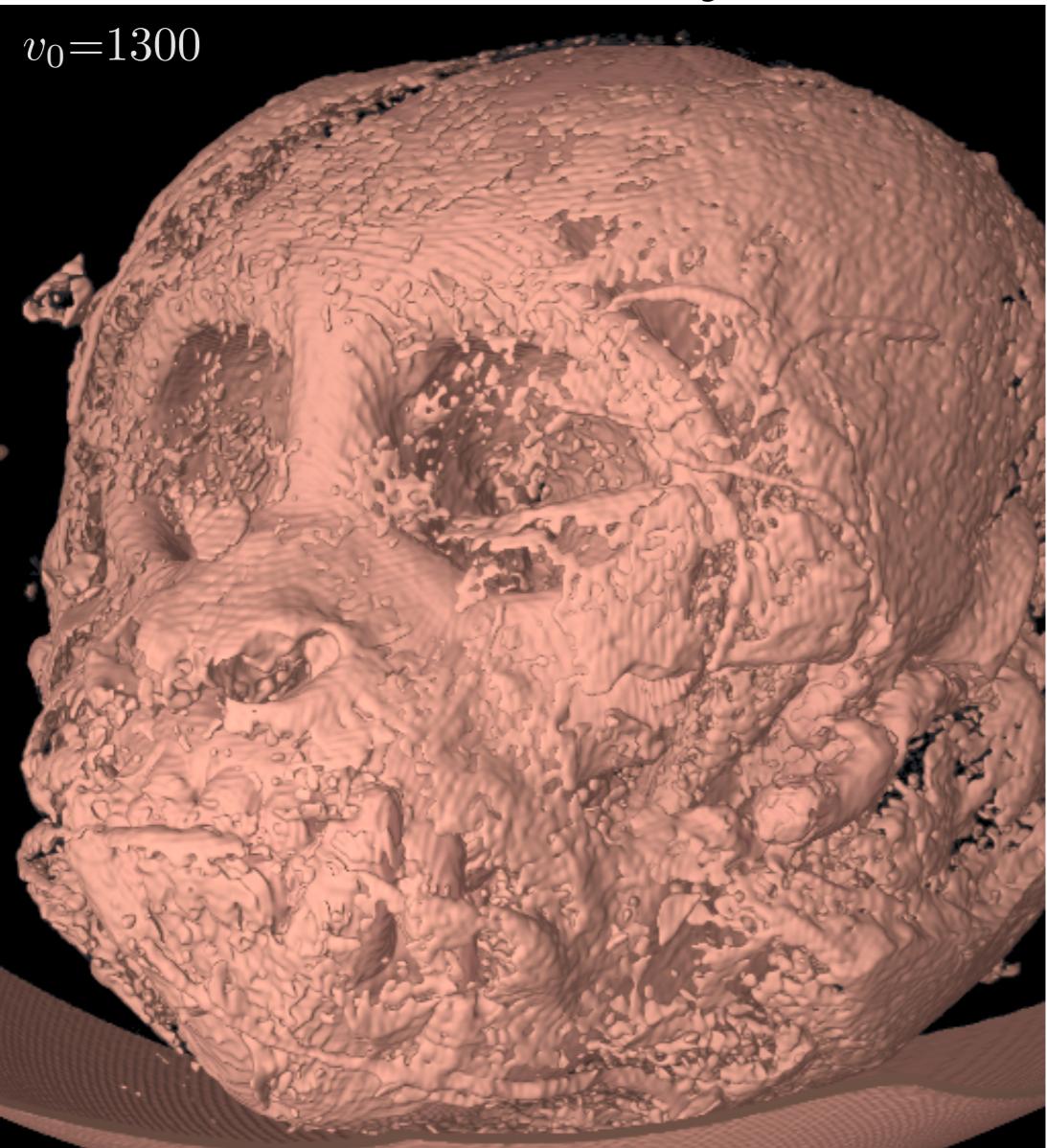
<http://diderot-language.cs.uchicago.edu/>

Acknowledgement: This material is based upon work supported by the NSF under Grants CCF-1446412 and CCF-1564298

Extra Slides

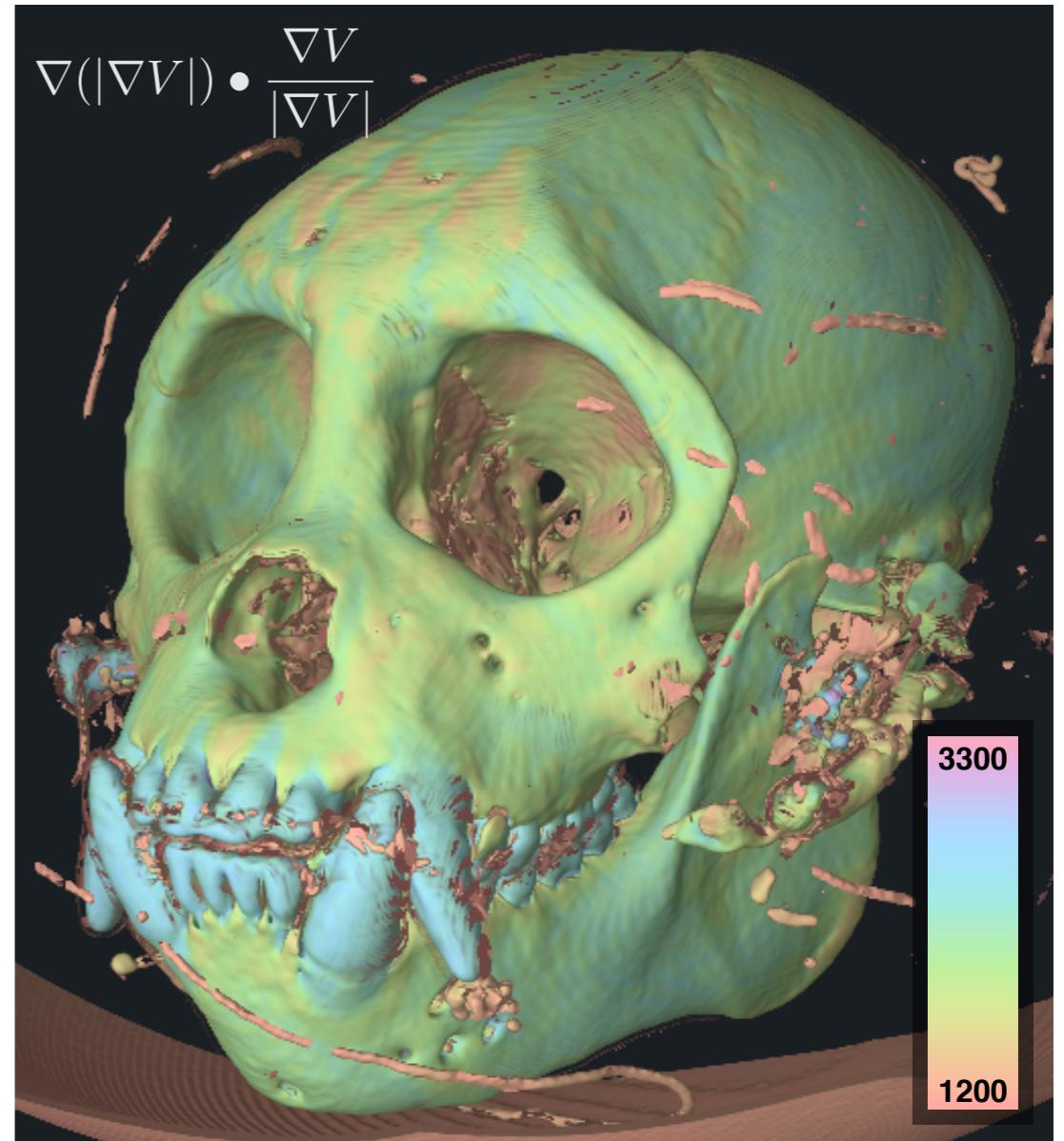
Results: Sophisticated Vis

Previously:



Iso-Surface

Now:



Canny Edges

Data provided by Callum Ross (U of Chicago)

Index based optimizations

Rewrites allow us to eliminate summations (i.e., effectively reduce the loop-nesting depth)

$$\nabla \bullet \nabla \times F \Rightarrow 0$$

$$\sum_{jk} \mathcal{E}_{ijk} \left(\frac{\partial}{\partial x_{ij}} \diamond e \right) \Rightarrow 0$$

$$\text{Trace}(u \otimes v) \rightarrow u \bullet v$$

$$(a \times b) \bullet (c \times d) \Rightarrow (a \bullet c)(b \bullet d) - (a \bullet d)(b \bullet c)$$

Compilation Issue

$$y = \lambda F, n \langle F_{ij}(n)F_{1k}(n) - F_{0i}(n)F_{jk}(n) + F_{ij}(n)F_{k1}(n) \rangle_{ijk}(\text{f}, n)$$

Compilation Issue

Field Terms

$$y = \lambda F, n \langle \underline{F_{ij}(n)} \underline{F_{1k}(n)} - \underline{F_{0i}(n)} \underline{F_{jk}(n)} + \underline{F_{ij}(n)} \underline{F_{k1}(n)} \rangle_{ijk} (f, n)$$

Compilation Issue

Field Terms

$$y = \lambda F, n \langle \underline{F_{ij}(n)} \underline{F_{1k}(n)} - \underline{F_{0i}(n)} \underline{F_{jk}(n)} + \underline{F_{ij}(n)} \underline{F_{k1}(n)} \rangle_{ijk} (f, n)$$

Compilation Issue

Field Terms

$$y = \lambda F, n \langle F_{ij}(n) \underline{F_{1k}(n)} - \underline{F_{0i}(n)} \underline{F_{jk}(n)} + F_{ij}(n) \underline{F_{k1}(n)} \rangle_{ijk}(\mathbf{f}, \mathbf{n})$$

Compilation Issue

Field Terms

$$y = \lambda F, n \left\langle \underline{F_{ij}(n)} \underline{F_{1k}(n)} - \underline{F_{0i}(n)} \underline{F_{jk}(n)} + \underline{F_{ij}(n)} \underline{F_{k1}(n)} \right\rangle_{ijk} (f, n)$$

Split

$$v_1 = \lambda F, n \langle F_{ij}(n) \rangle_{ij} (f, n)$$

$$v_2 = \lambda F, n \langle F_{1i}(n) \rangle_i (f, n)$$

$$v_3 = \lambda F, n \langle F_{0i}(n) \rangle_i (f, n)$$

$$v_4 = \lambda F, n \langle F_{ij}(n) \rangle_{ij} (f, n)$$

$$v_5 = v_1$$

$$v_6 = \lambda F, n \langle F_{i1}(n) \rangle_i (f, n)$$

Compilation Issue

Field Terms

$$y = \lambda F, n \left\langle \underline{F_{ij}(n)} \underline{F_{1k}(n)} - \underline{F_{0i}(n)} \underline{F_{jk}(n)} + \underline{F_{ij}(n)} \underline{F_{k1}(n)} \right\rangle_{ijk} (f, n)$$

Split

$$v_1 = \lambda F, n \left\langle F_{ij}(n) \right\rangle_{ij} (f, n)$$

$$v_2 = \lambda F, n \left\langle F_{1i}(n) \right\rangle_i (f, n)$$

$$v_3 = \lambda F, n \left\langle F_{0i}(n) \right\rangle_i (f, n)$$

$$v_4 = \lambda F, n \left\langle F_{ij}(n) \right\rangle_{ij} (f, n)$$

$$v_5 = v_1$$

$$v_6 = \lambda F, n \left\langle F_{i1}(n) \right\rangle_i (f, n)$$

Compilation Issue

Field Terms

$$y = \lambda F, n \left\langle \underline{F_{ij}(n)} \underline{F_{1k}(n)} - \underline{F_{0i}(n)} \underline{F_{jk}(n)} + F_{ij}(n) \underline{F_{k1}(n)} \right\rangle_{ijk} (f, n)$$

Split

$$v_1 = \lambda F, n \langle F_{ij}(n) \rangle_{ij} (f, n)$$

$$v_2 = \lambda F, n \langle F_{1i}(n) \rangle_i (f, n)$$

$$v_3 = \lambda F, n \langle F_{0i}(n) \rangle_i (f, n)$$

$$v_4 = v_1$$

$$v_5 = v_1$$

$$v_6 = \lambda F, n \langle F_{i1}(n) \rangle_i (f, n)$$

Compilation Issue

Field Terms

$$y = \lambda F, n \left\langle \underline{F_{ij}(n)} \underline{F_{1k}(n)} - \underline{F_{0i}(n)} \underline{F_{jk}(n)} + F_{ij}(n) \underline{F_{k1}(n)} \right\rangle_{ijk} (f, n)$$

Split

$$v_1 = \lambda F, n \langle F_{ij}(n) \rangle_{ij} (f, n)$$

$$v_2 = \lambda F, n \langle F_{1i}(n) \rangle_i (f, n)$$

$$v_3 = \lambda F, n \langle F_{0i}(n) \rangle_i (f, n)$$

$$v_4 = v_1$$

$$v_5 = v_1$$

$$v_6 = \lambda F, n \langle F_{i1}(n) \rangle_i (f, n)$$

Slice

$$v_1 = \lambda F, n \langle F_{ij}(n) \rangle_{ij} (f, n)$$

$$v_2 \rightarrow \lambda T \langle T_{1i} \rangle_i (t)$$

$$v_3 \rightarrow \lambda T \langle T_{0i} \rangle_i (t)$$

$$v_4 = v_1$$

$$v_5 = v_1$$

$$v_6 \rightarrow \lambda T \langle T_{i1} \rangle_i (t)$$

Compilation Issue

Field Terms

$$y = \lambda F, n \left\langle \underline{F_{ij}(n)} F_{1k}(n) - F_{0i}(n) F_{jk}(n) + F_{ij}(n) F_{k1}(n) \right\rangle_{ijk} (f, n)$$

Split

$$v_1 = \lambda F, n \langle F_{ij}(n) \rangle_{ij} (f, n)$$

$$v_2 = \lambda F, n \langle F_{1i}(n) \rangle_i (f, n)$$

$$v_3 = \lambda F, n \langle F_{0i}(n) \rangle_i (f, n)$$

$$v_4 = v_1$$

$$v_5 = v_1$$

$$v_6 = \lambda F, n \langle F_{i1}(n) \rangle_i (f, n)$$

Slice

$$v_1 = \lambda F, n \langle F_{ij}(n) \rangle_{ij} (f, n)$$

$$v_2 \rightarrow \lambda T \langle T_{1i} \rangle_i (t)$$

$$v_3 \rightarrow \lambda T \langle T_{0i} \rangle_i (t)$$

$$v_4 = v_1$$

$$v_5 = v_1$$

$$v_6 \rightarrow \lambda T \langle T_{i1} \rangle_i (t)$$

Compilation Issue

Field Terms

$$y = \lambda F, n \left\langle F_{ij}(n) F_{1k}(n) - F_{0i}(n) F_{jk}(n) + F_{ij}(n) F_{k1}(n) \right\rangle_{ijk} (f, n)$$

Split

$$v_1 = \lambda F, n \langle F_{ij}(n) \rangle_{ij} (f, n)$$

$$v_2 = \lambda F, n \langle F_{1i}(n) \rangle_i (f, n)$$

$$v_3 = \lambda F, n \langle F_{0i}(n) \rangle_i (f, n)$$

$$v_4 = v_1$$

$$v_5 = v_1$$

$$v_6 = \lambda F, n \langle F_{i1}(n) \rangle_i (f, n)$$

Slice

$$v_1 = \lambda F, n \langle F_{ij}(n) \rangle_{ij} (f, n)$$

$$v_2 \rightarrow \lambda T \langle T_{1i} \rangle_i (t)$$

$$v_3 \rightarrow \lambda T \langle T_{0i} \rangle_i (t)$$

$$v_4 = v_1$$

$$v_5 = v_1$$

$$v_6 \rightarrow \lambda T \langle T_{i1} \rangle_i (t)$$

$$y = \lambda T, B, C, D \langle T_{ij}B_k - C_iT_{jk} + T_{ij}D_k \rangle_{ijk} (v_1, v_2, v_3, v_6)$$