Underwater holograms, plankton, and the Grid

J.J. Nebrensky and P.R. Hobson

School of Engineering and Design
Brunel University, UK

BRUNEL UNIVERSITY
WEST LONDON

Topics

- *Holography underwater… WHY???*
- *Plankton basics*

- Digital Holography
- What is the Grid?
- How can we use the Grid to do it better?

The need to study marine organisms

- **Behaviour of marine biological communities**
  - plays important role in understanding global environment
    - e.g. atmospheric interactions {global warming}
    - e.g. biological interactions {fisheries, algal blooms}
- **Modelling of chemical cycles** assisted by
  - study of aggregates of biotic or abiotic particles
  - knowledge of distribution & interaction dynamics
    - the **inter-relationships** between various organisms
- **Accuracy of measurements limited**
  - absence of good measurement techniques
    - counters, sampling bottles or photography
  - frailty & wide size range & complexity of aggregates
  - particles vary in size from sub-micron to several millimetres

The benefits of holography in marine biology

- **Records live species in natural environment**
  - non-intrusive, non-destructive, *in situ* interrogation
  - can record large volumes of water column in one short exposure
- **True three-dimensional imaging of organisms**
  - retention of parallax & perspective information
  - high image resolution over large depth-of-field
  - wide recording dynamic range
- **Ability to isolate individual planes of the image**
  - move viewing plane through image volume to bring individual species into focus
- **Aids study of marine biological communities**
  - species identification & classification at genus level
  - measurement of size & relative position of individual organisms
  - measurement of distribution of organisms & interrelationships

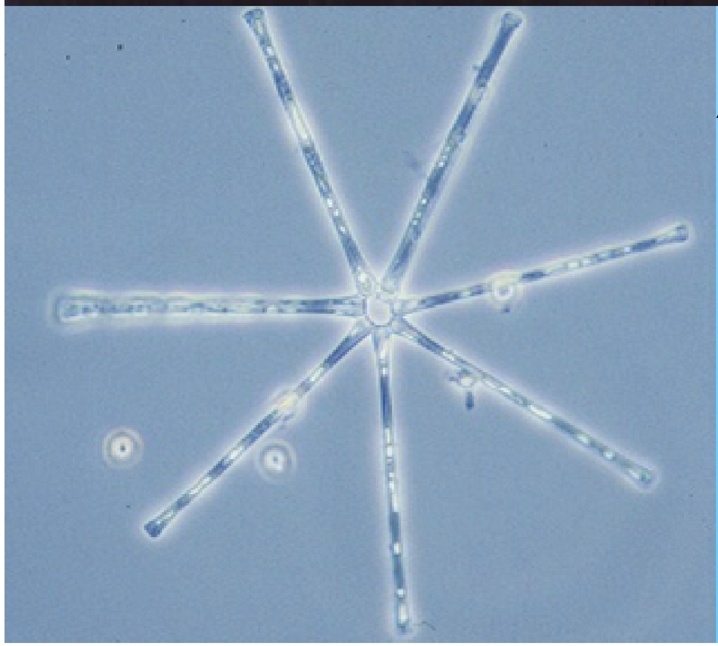
HOLOMAR

Plankton Basics

NERC news

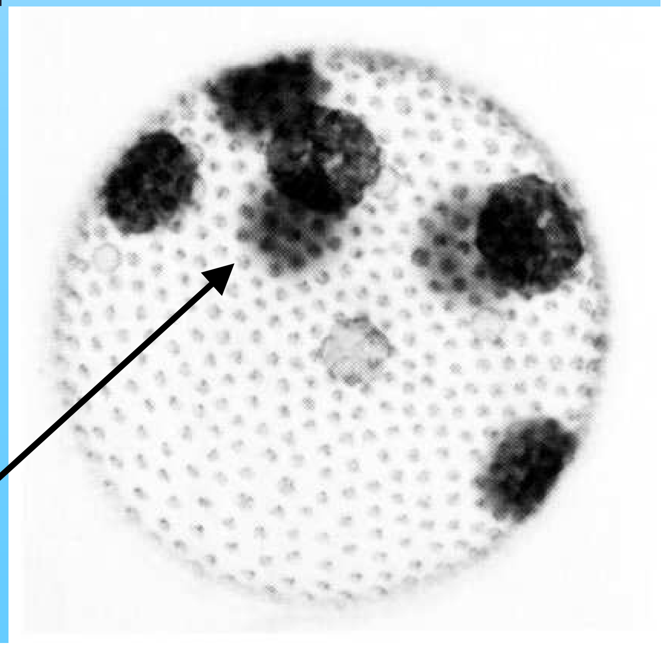
Fish eat zooplankton, e.g. *Calanus*, which in turn eat (smaller) phytoplankton

***Volvox Aereus* – one colony (~200 µm dia.) can be 100s of individuals**



***Asterionella Formosa* – a freshwater species that forms small groups (1-10 animals, each ~40 µm long.**

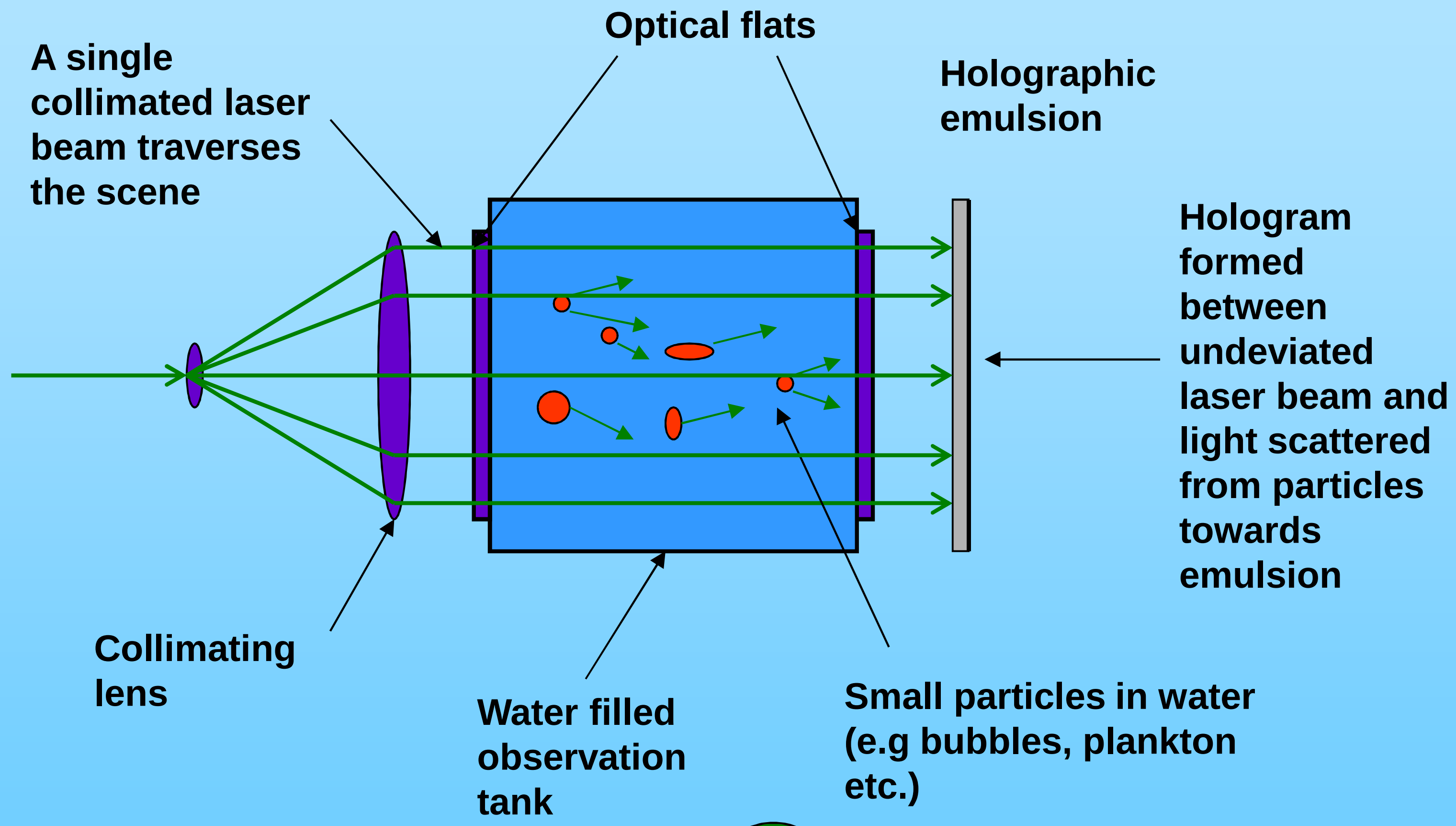
Baby colonies



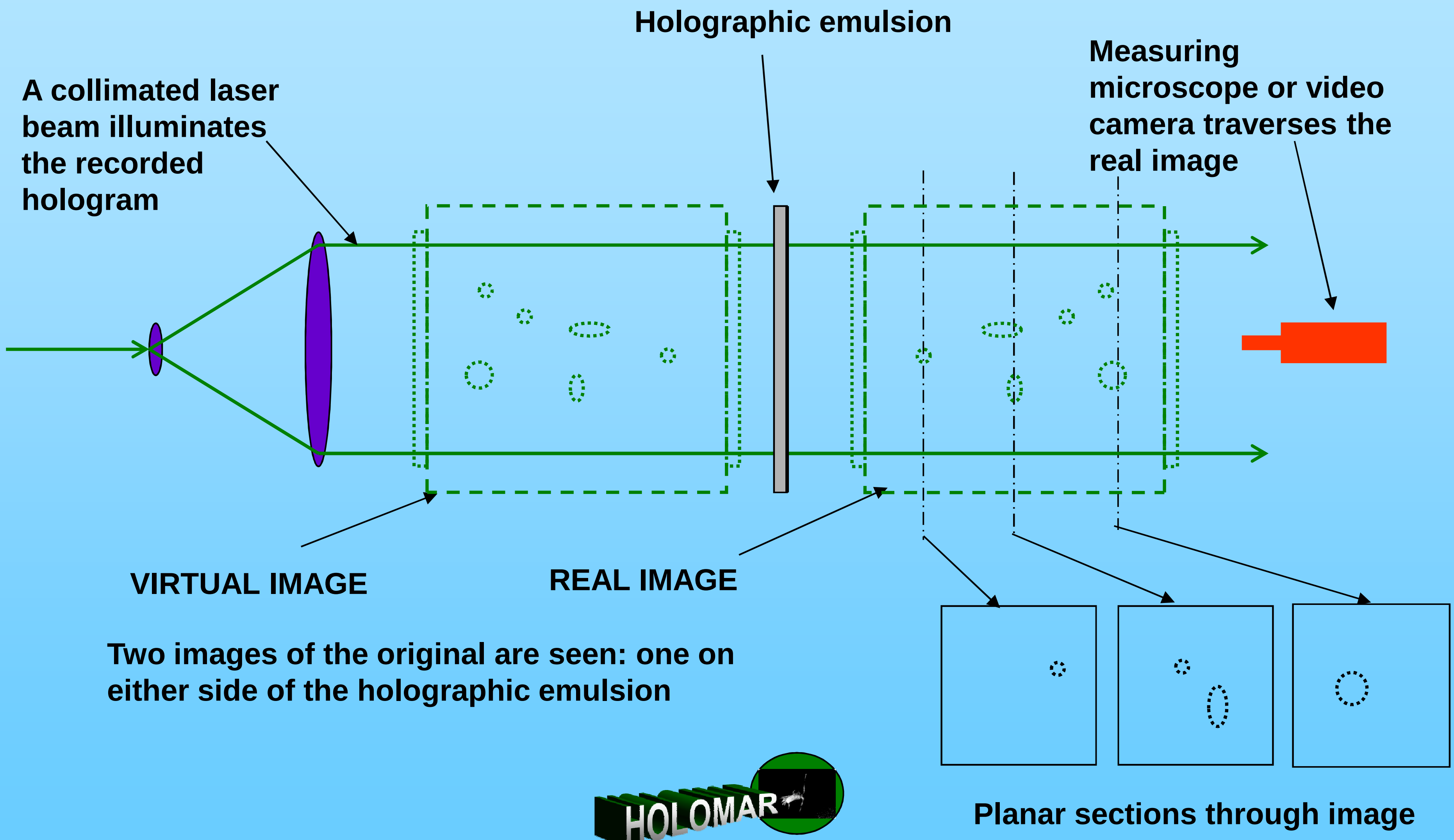
via R. Lampitt

H. Volten – Doctoral Thesis

Recording an In-line hologram



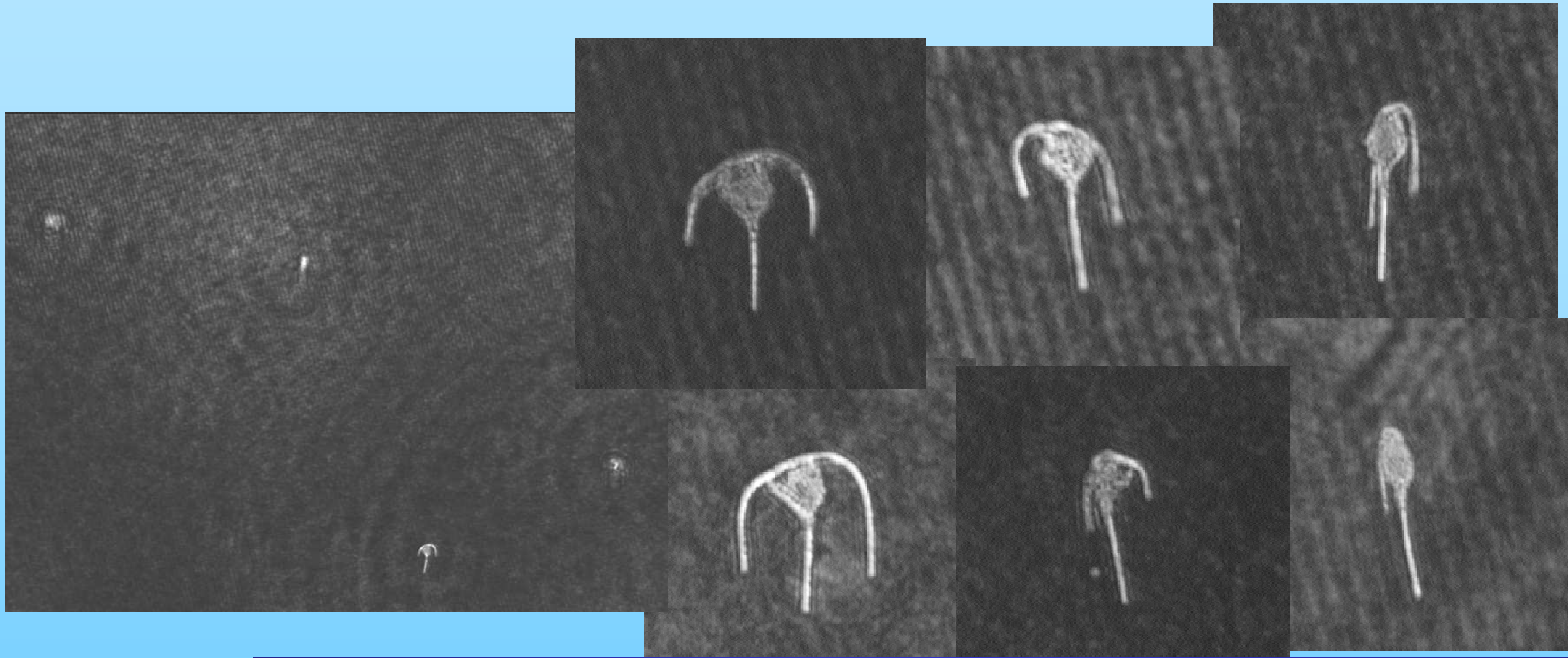
Replaying an in-line hologram



Two images of the original are seen: one on either side of the holographic emulsion

HOLOMAR

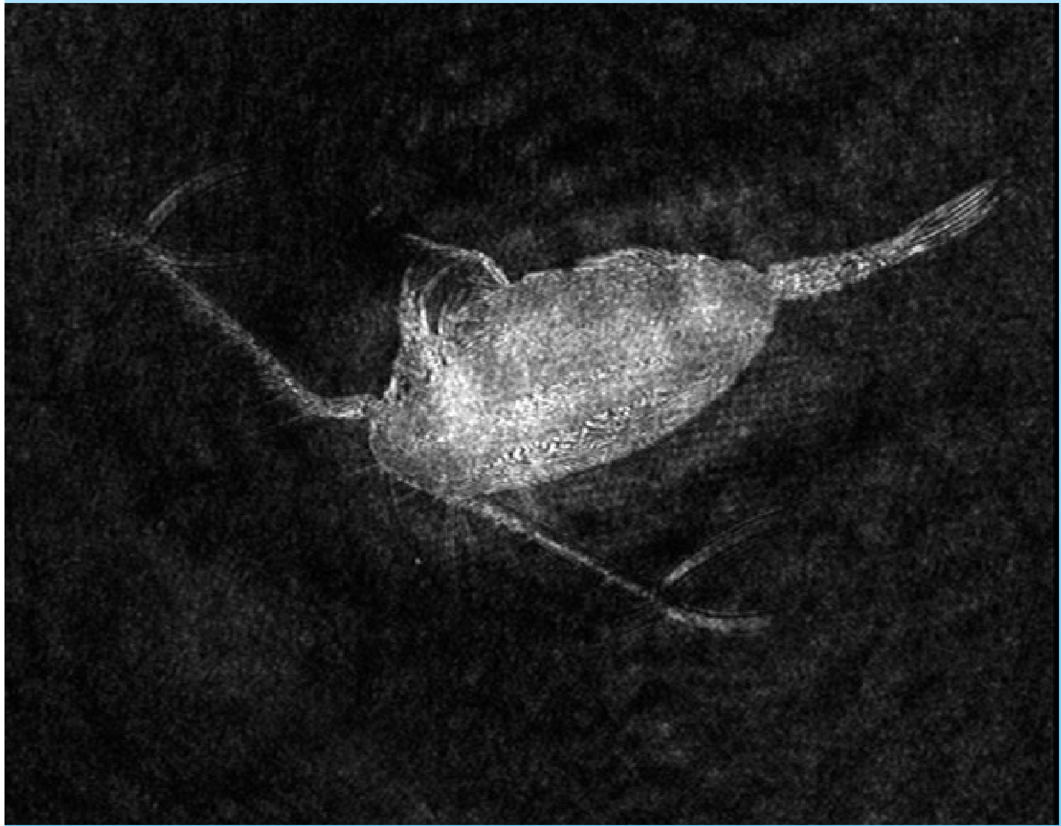
A selection of *Ceratium* images from a single in-line hologram

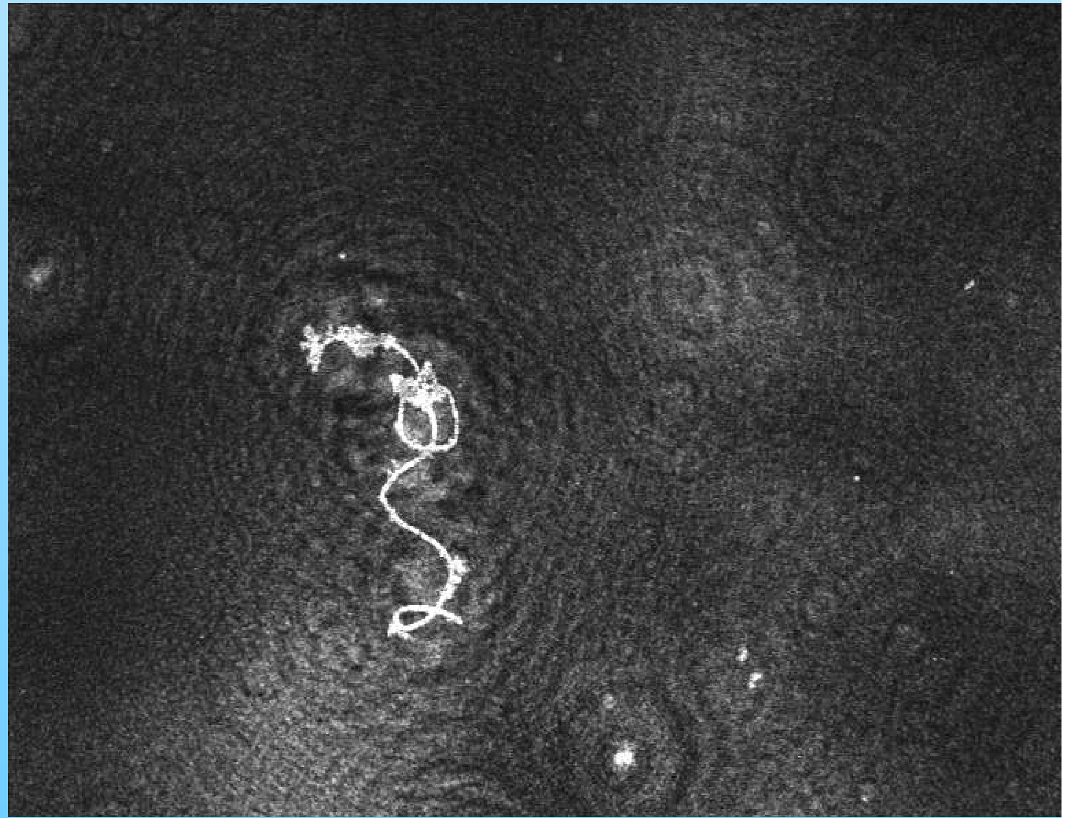


The image on the left is a 5.3 mm by 4 mm plane of an in-line hologram at low magnification. The other images, at higher magnification, show a selection of *Ceratium* organisms (approximately **200 μm** long) from the same hologram. They were all recorded by in-line holography in a single exposure of a 2.5 litre volume of water. Each organism was at a different location.

HOLOMAR

Loch Etive plankton – in-line





The Holographic Data Problem

If a photograph is worth a thousand words, then a hologram is worth a million photographs

(*after Keith Pennington)*

- Our in-line sample volume is 100 mm diameter and 0.5 m long
- The *low*-magnification camera has a field of view about 9 by 6 mm
- That’s around 200 images for each slice
- Even assuming a slice every 0.1 mm (bigger than many objects)
  - **A million images in each holographic plate**

The Holographic Data Problem

Manual or semi-automatic analysis generally takes about a person-week for each holographic exposure

- Brown (ice crystals in clouds) - “a few hours” for a 150 cm^3 sample volume
- Borrmann & Jaenicke (snow / fog) - 32 hours / hologram: 8 cm^3 and 1000 objects
- Vössing *et al*. (snow / fog) - up to 70 hours for 1 hologram of a 500*l* volume
- Katz *et al*. (plankton) - two weeks for each hologram of 300 to 2000 cm^3

The main issue is operator fatigue leading to measurement errors

Automated analysis is needed ...

The Holographic Data Problem

... but still a challenge

At *high* magnification (a 1 mm by 0.7 mm view), one

plate can generate **30 *Tera*Bytes** of raw data

- **Need to extract and visualise information, not data**
- **How does one characterise the 3-d, projected real image ?**
  - **e.g. brightness and contrast: how to find the brightest and darkest voxels in that 30 Tb?**
  - **Real image properties both fixed in plate and depend on replay laser and viewing camera**

HOLOMAR

Image extraction and processing

Steps:

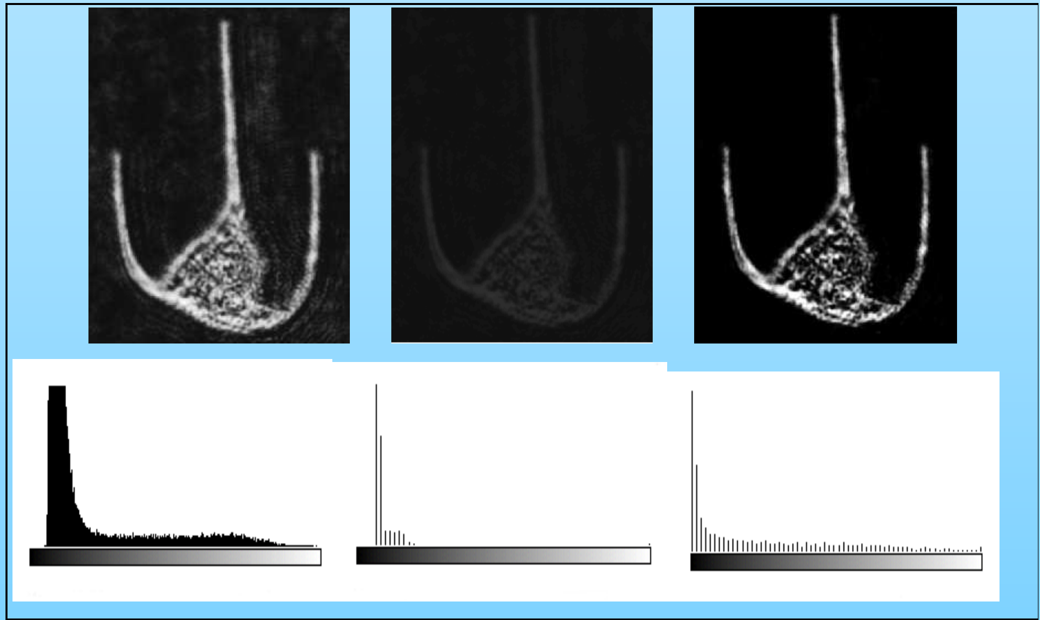
- **Global adjustment of hologram for brightest and sharpest image**
  - orientation of plate holder and angle of reference beam
- **Scan videocamera through depth; capture successive images**
- **Digital processing for image enhancement**
  - cleaning and background removal
  - object tracking
  - best focus
  - image enhancement
  - segmentation
- **Species identification**
  - based on neural networks recognition

Results:

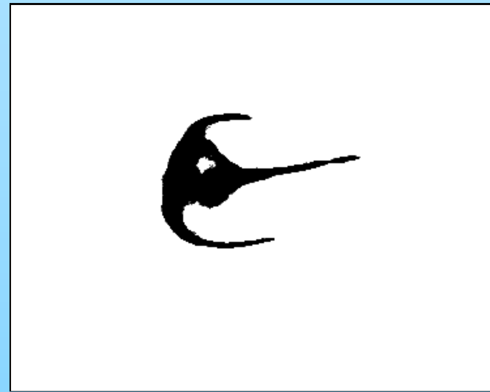
- **Size measurement & relative position**
- **Measurement of local concentration and distribution by category**

HOLOMAR

Image enhancement & species identification



Original, cropped & stretched images
with relative frequency distributions



Filtered, segmented & pre-processed image prior to identification by neural net techniques

Object classification - neural net

- **7 Inputs:**
  - Blob size
  - Elongation
  - first five Hu moments
- **7 neuron hidden layer**
- **5 Outputs - in-line:**
  - 3 types of phytoplankton
  - Zooplankton
  - Floc (dead or inorganic debris)

Class	Well classified (%)
Asterionella	95.2
Ceratium	94.4
Thalassiosira	85.7
Zooplankton	100
Floc	100

HOLOMAR

Digital Holography

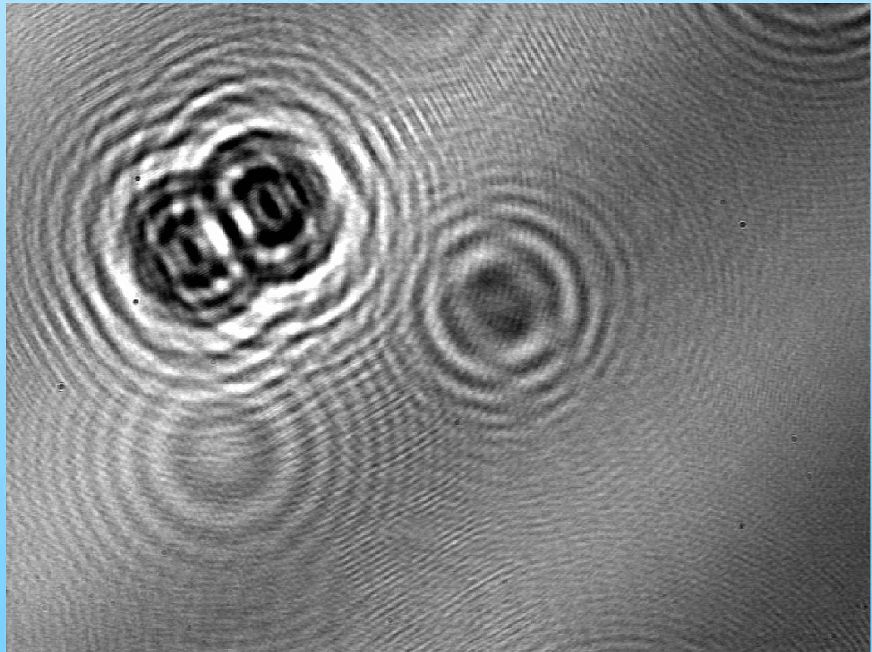
Instead of using photographic film, it is possible to capture the hologram directly on to the CCD, and then reconstruct numerically within a computer.

This avoids the need to handle glass plates within the holocamera, eliminates chemical development and can compensate for the refractive index change.

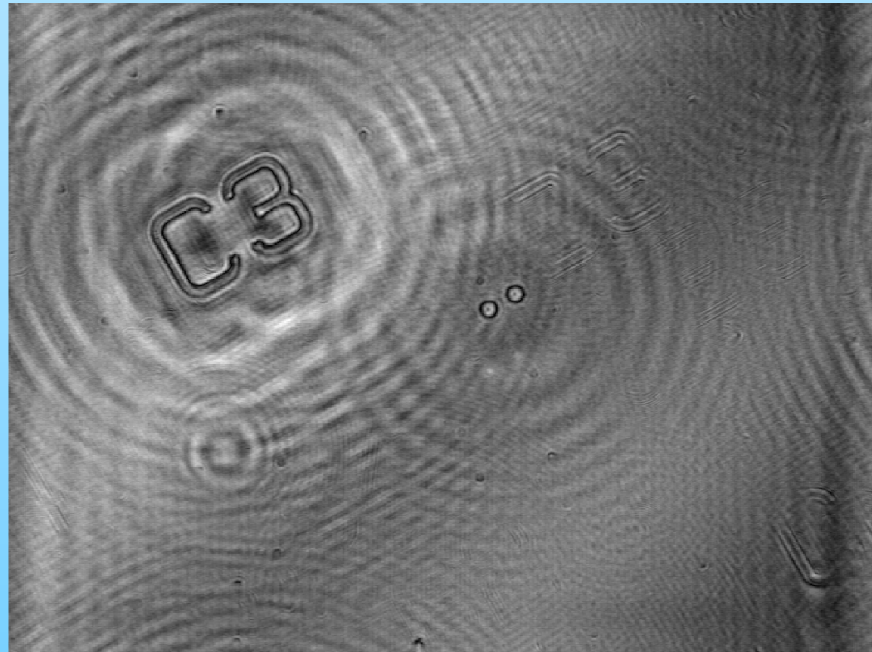
- **Holocamera smaller (less intrusive) and lighter**
- **Don’t need to return to shore before processing**
- **Results already in digital form for further analysis and visualisation**

Digital Holography

Sample results obtained from an array of pairs of opaque discs by Marc Fournier-Carrié (Brunel University):



An in-line hologram of a test target, captured from a CCIR videocamera



The in-focus objects, regenerated within the computer

Digital holosub results

Nauplius copepod goes after a dinoflagellate, but it escapes!

(Video removed – see website below for details)

Katz / JHU website **http://www.me.jhu.edu/~lefd/shc/HoloSub_site/Digital_Holosub.htm**

Digital Holography

To find objects for further analysis, we must reconstruct a series of slices and then search through them. Numerical reconstruction is computationally heavy – multiple FFTs – and intermediate data volumes are huge.

We have our own FFTW-based reconstruction software “HoloPlay” for single image planes from in-line holograms. This has been tested with images from an Atmel Camelia camera (8 Mpixel, 12-bit depth) †.

The same source code compiles and runs both on Visual C++ 6 on Windows 2000, and on GCC 3.2 on Linux (RedHat 7.3 / SLC 3).

† based on UG project work by Paul Fryer

How heavy?

Timings based on original reconstruction software:

```
f: P2 400 MHz 384 Mb (RH6.2)

SAVVAS1.PGM   768x576 8bit DFT with no padding  (0.44Mpixels)    914 s
SAVVAS1.PGM   768x576 8bit DFT pad to 1024x1024 (   1Mpixels)   3305 s
SAVVAS1.PGM   768x576 8bit FFT pad to 1024x1024 (   1Mpixels)     17 s
8MC.PGM      1024x512 8bit FFT with no padding  ( 0.5Mpixels)      8 s
8MC.PGM      1024x512 8bit DFT with no padding  ( 0.5Mpixels)   1240 s
8MWIRE.PGM  2300x3500 8bit FFT pad to 4096x4096 (  16Mpixels)    671 s
8MWIRE.PGM  2300x3500 8bit DFT pad to 4096x4096 (  16Mpixels) 216504 s
8MWIRE.PGM  2300x3500 8bit DFT with no padding  (   8Mpixels)  73272 s
```

Image enhancement & species identification:

```
   P2 400MHz 128Mb (NT 4)
42 mins HS20120A ci6a 768x566 1-459 no objects
6'04    HS20120A co2a 768x574 1-61  no objects
```
(<10 seconds per frame)

How heavy?

For each CCIR frame, object tracking and extraction is quicker than the reconstruction.
(We don't yet know how this scales with frame size)

Timings based on original reconstruction software:

```
g: Athlon XP2600+ (1900MHz) 1024Mb (RH7.3)
SAVVAS1.PGM   768x576 8bit FFT pad to 1024x1024 (   1Mpixels)      3 s
8MC.PGM      1024x512 8bit FFT with no padding  ( 0.5Mpixels)      2 s
8MWIRE.PGM  2300x3500 8bit FFT pad to 4096x4096 (  16Mpixels)    101 s
```

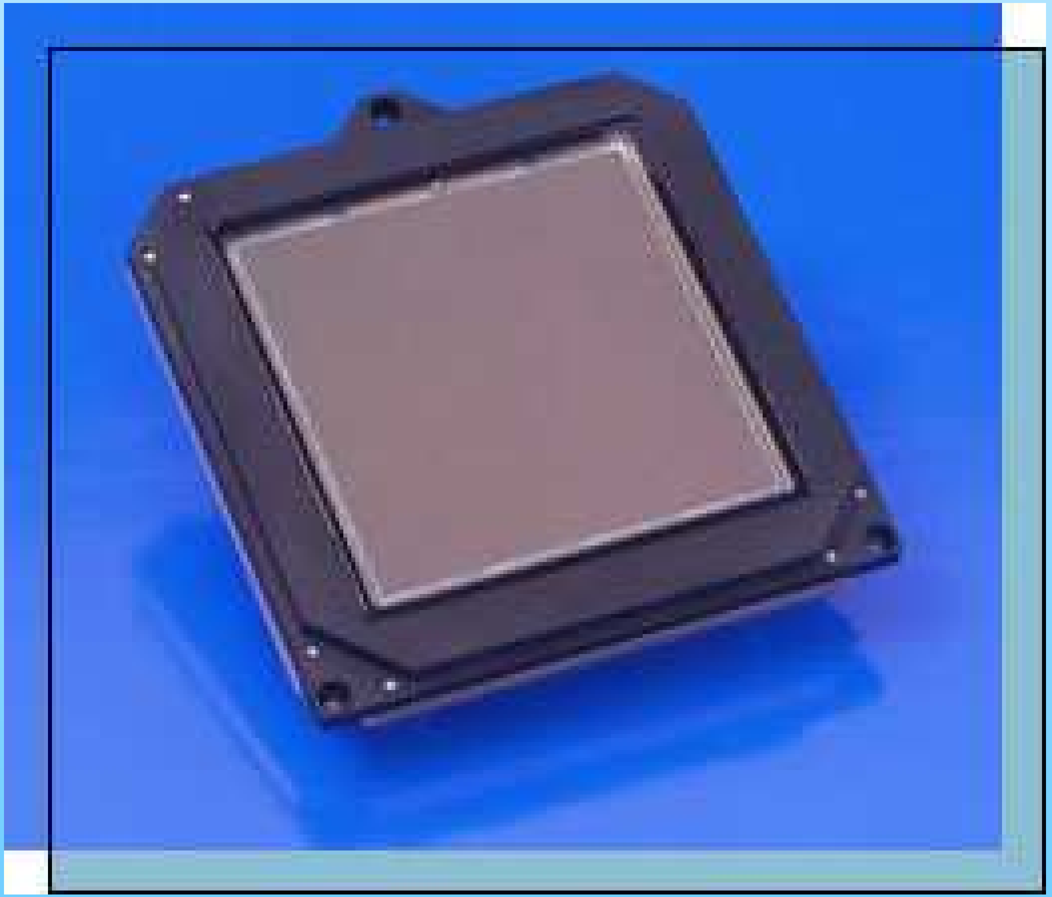
(but object tracking and extraction is quicker, too!)

About 2 mins for each plane – *many* planes (100's) in a single hologram!

How heavy!

Fairchild Imaging CCD595:

- **9216 x 9216 full frame CCD**
- **8.75 µm x 8.75 µm pixels**
- **80 mm x 80 mm image area**
- **100% fill factor**
- **€50-100 000**



85 Mpixels!

http://www.fairchildimaging.com/main/area_595.htm

What about Grid computing?

- Big buzz, generally about big number-crunching and huge data storage… J

- Users identified via X.509 certificates:
  must be in a suitable VO, then single sign-on
  F now on LTWO

Constantly changing acronyms - terminology here is a personal mish-mash L

So:
The Grid promises to let you provide serious computing resources to your RAs as readily as mains electricity…

Grid Computing in a Nutshell

BBC Cult (bbc.co.uk)

The venerable "Electrical Power" analogy

How can we use the Grid to do it better?

Use Grid to reconstruct the planes in parallel:

- **store digital hologram (and binary) at SE**
- **submit control file for each plane (`edg-job-submit`) and monitor with LCG GUI monitor (`edg-wl-ui-gui-jobmonitor`)**
- **recover image planes to UI via local SE**
- **use existing tracking and identification code**

Need to keep track of reconstructed planes (which Grid job produced the next image) – an extra layer of software.

Digital Holography on the Grid

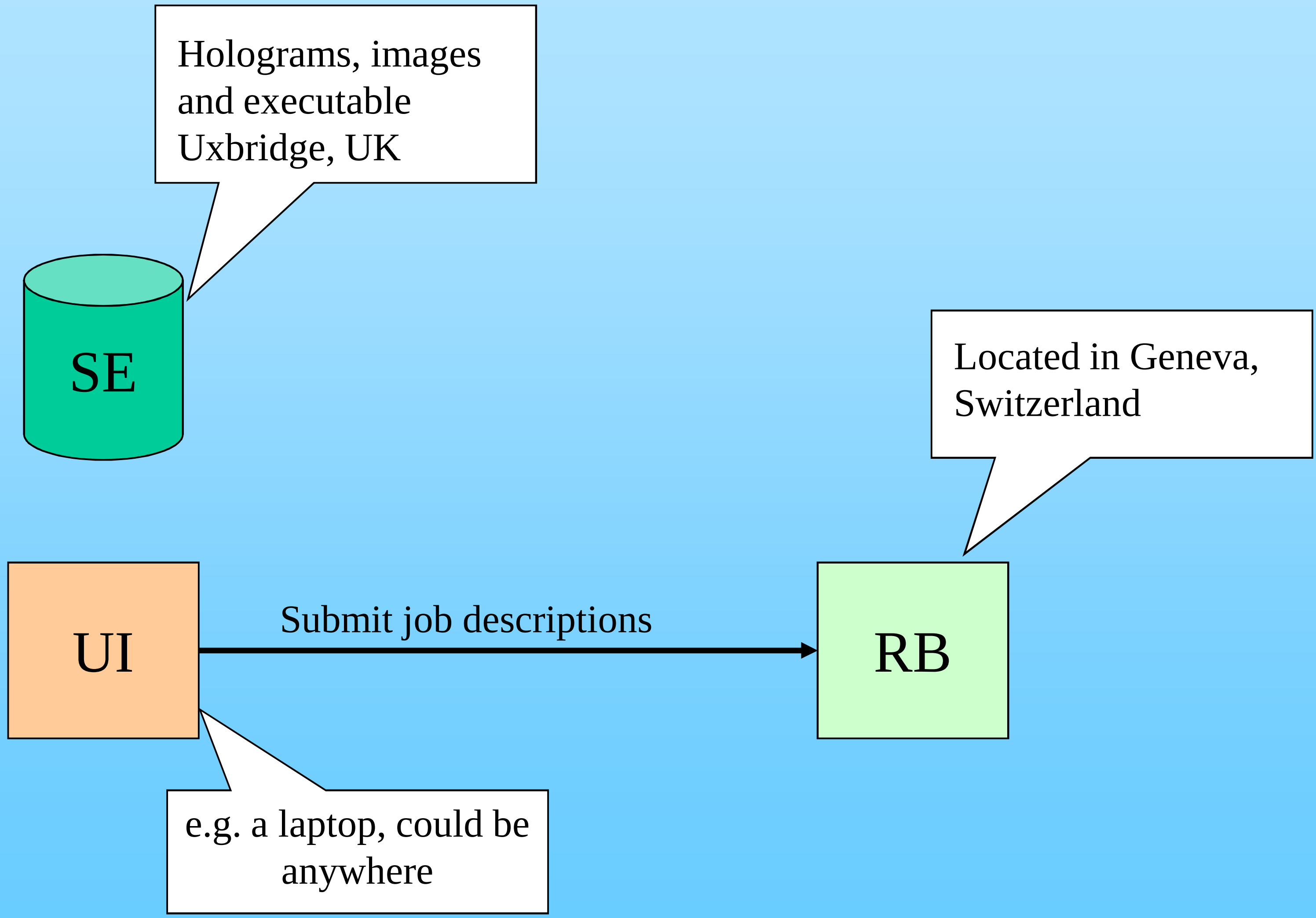
We have placed a sample hologram (and binary) on the SE associated with the BITLab facility at Brunel University.

HoloPlay was then run on resources around the Grid, with the reconstructed images being stored on the Brunel SE.

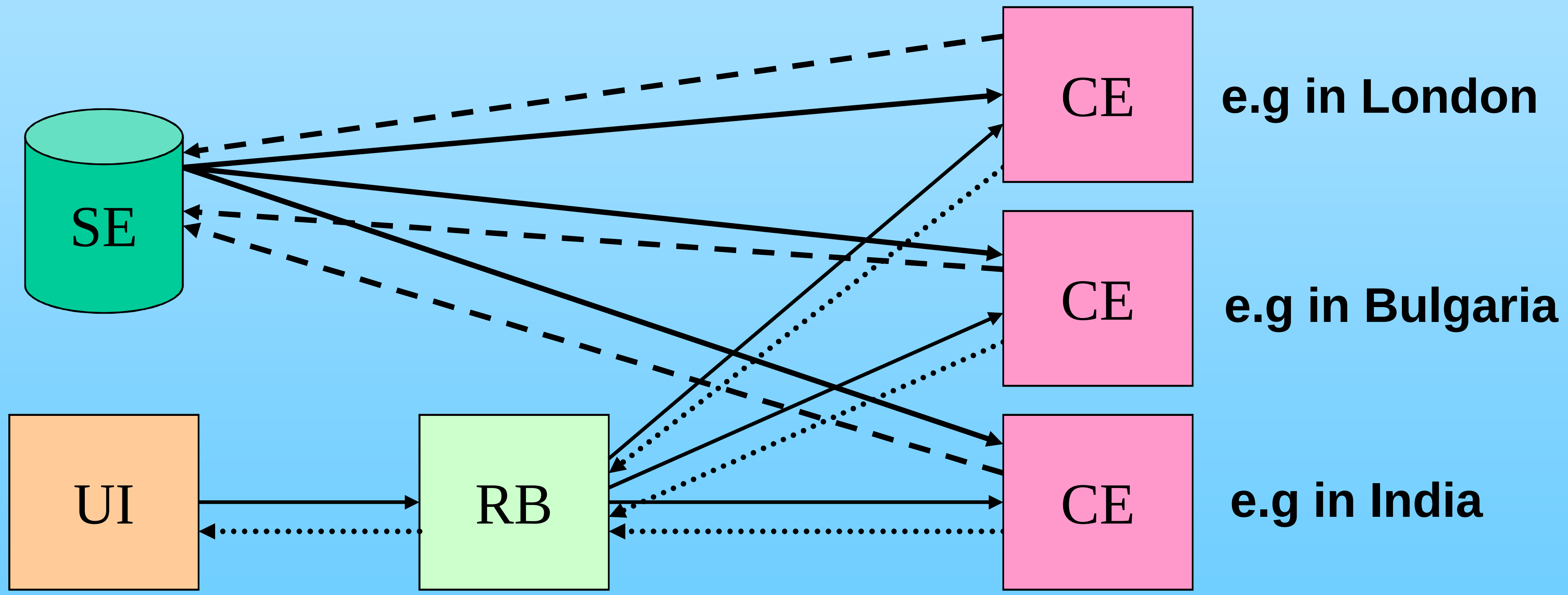
Our sample holograms are 2300 by 3500 pixels with 16-bit depth. We use PGM format image files, which can be up to 40Mb.

The hologram and images are thus gzipped before being uploaded to the SE.

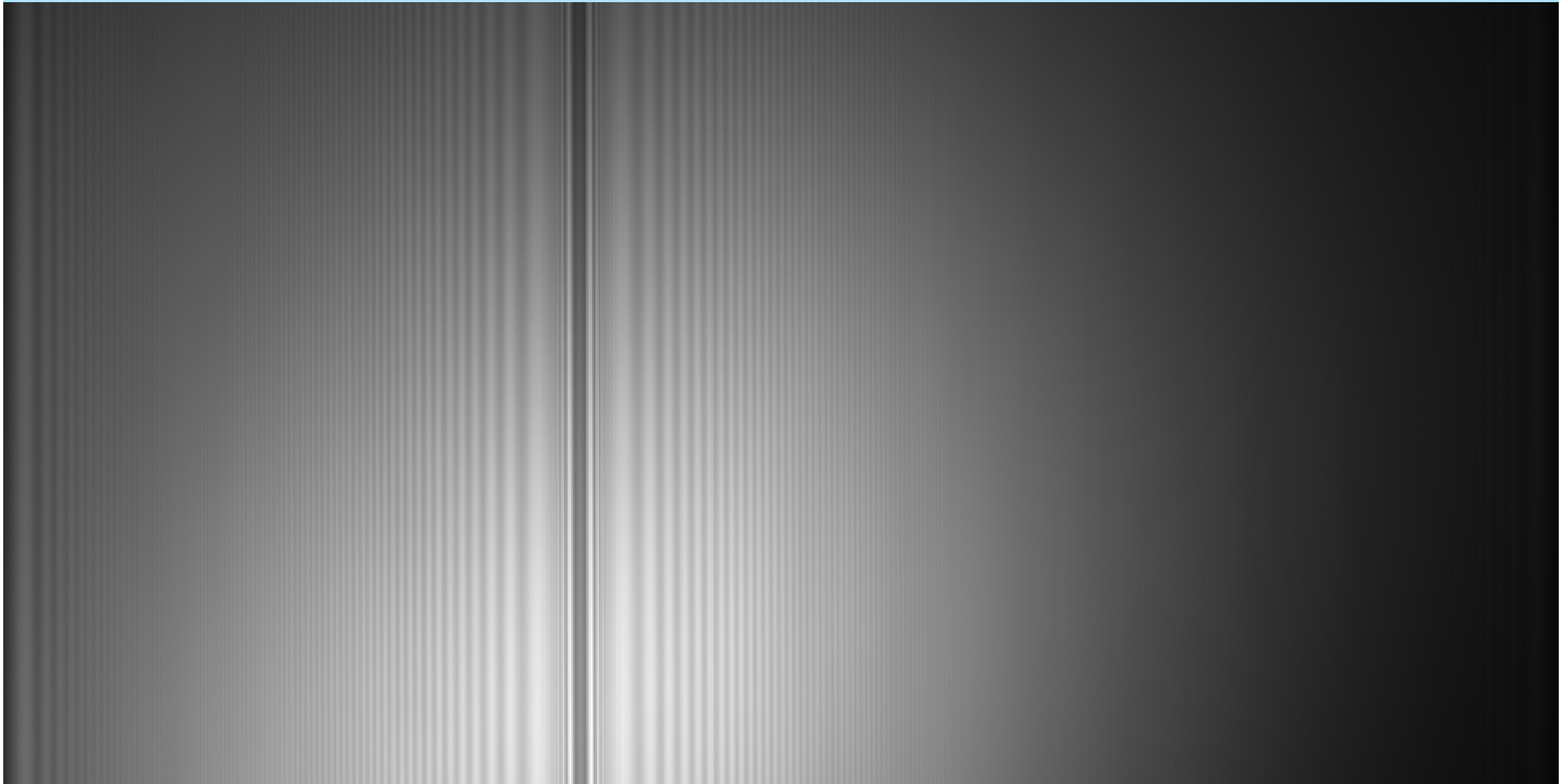
Some of the players



Reconstruction on the Grid



First known reconstruction using WAN Grid



An out-of-focus reconstruction of a 50 micron wire.

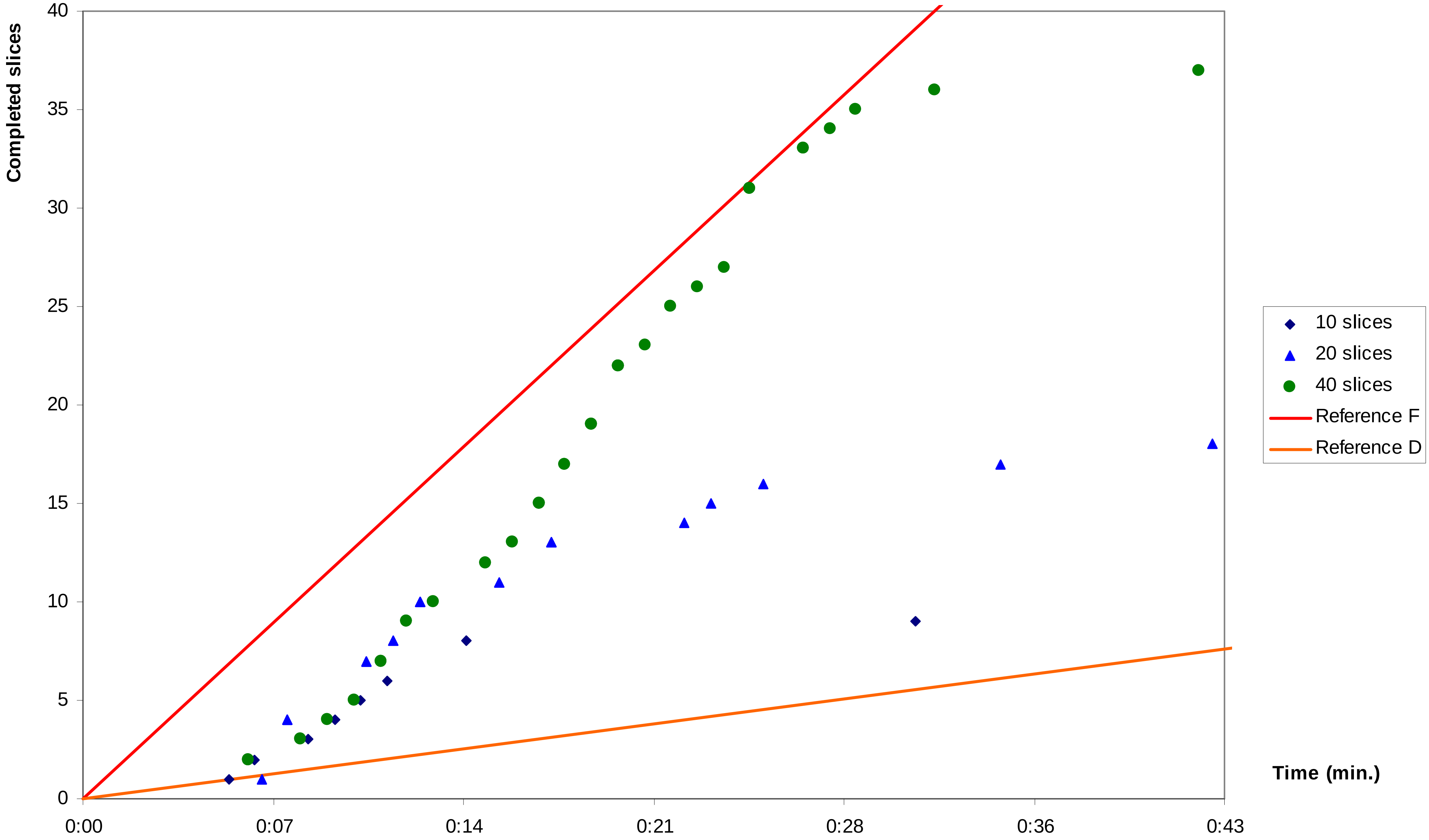
Evaluation

We’ve submitted batches of 10 – 40 single slices, and looked at how long it takes between starting the submission and the replayed images arriving at the SE.

Compare this with replaying the slices sequentially on a single machine (~1’15” each on the UI).

- Most of the volume is generated much faster than with the single CPU
- Some stragglers mean that it takes a bit longer to get all of the volume back than to compute it on a single PC.

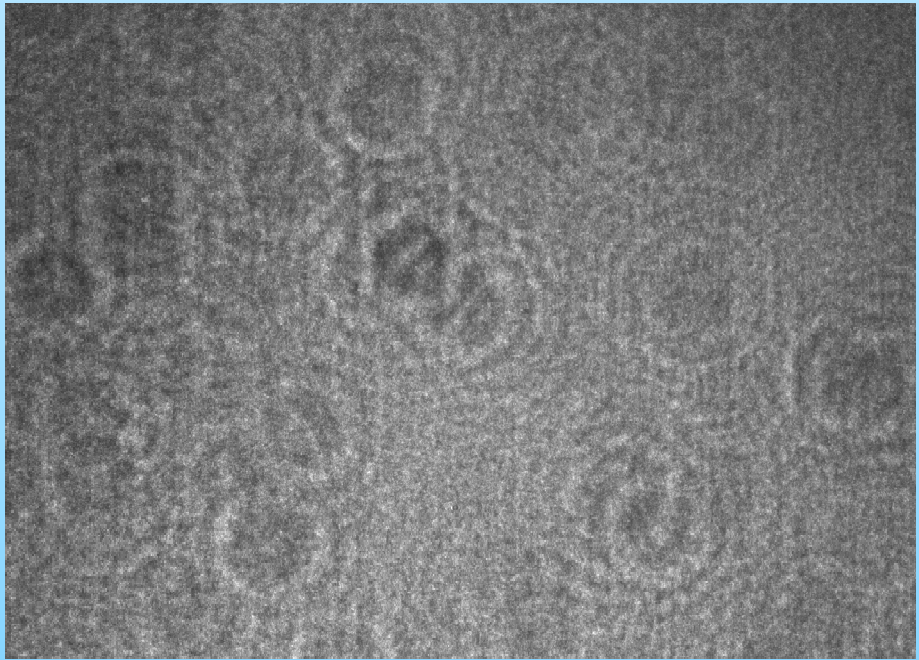
Job Progress

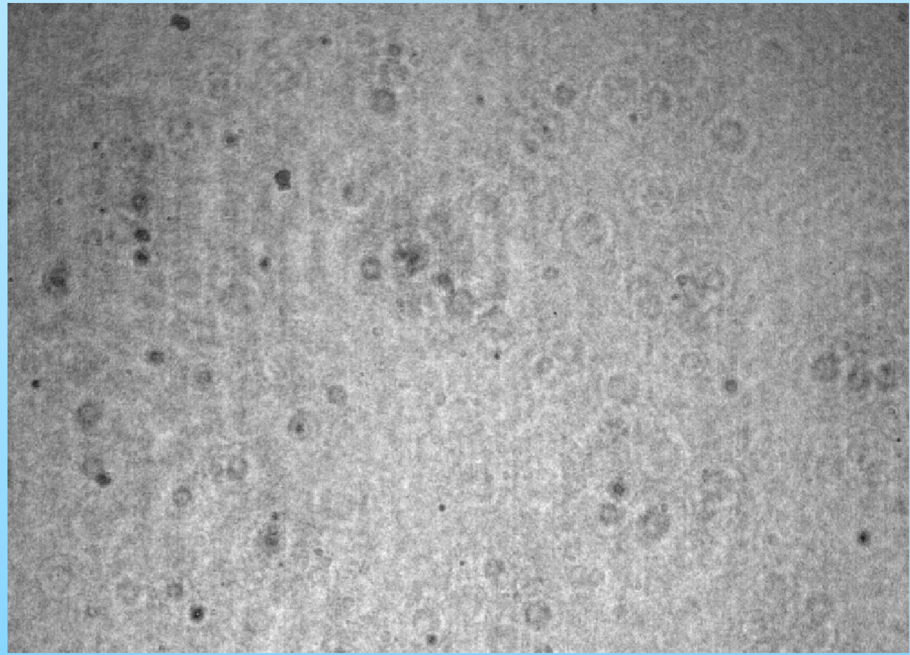


Problems/solutions

- No requirement in JDL that the WN used for replay must be comparable with our reference machine
- Overheads in job submission and data transfer similar to task itself. Each job doing several slices would be better.
- Network connectivity at remote sites – e.g. it took 20 mins to transfer a reconstructed 40Mb image back from India
- Data bandwidth into/from SE – if too many remote jobs simultaneously upload results it will slow right down
- Misconfigured sites and middleware problems mean that Grid job efficiency is about 90% - i.e. 10% of jobs either fail completely (never return data) or are resubmitted within the Grid infrastructure which can take >12 hours… Need something to track progress so that gaps can be filled in quicker.

What next?





- Currently looking at particles in water (“Fillite”)
- LT2 resources should (hopefully) have fairly homogeneous networking and response
- Look at improved submission strategies – more slices per Grid job? Return when complete or at end? Different start and finish SEs?

Related hologrammetry applications

High-resolution imaging & measurement in ‘dense’ media:

- **Offshore inspection**
  - archiving, corrosion pitting, damage, dimensional measurement
- **Nuclear fuel inspection**
- **Bubble chamber diagnostics -** analysis of nuclear particle tracks
- **Marine life, organisms, bubble fields**
  - recording / monitoring of coral reefs
  - sediment transport / tracking
  - cavitation nuclei
- **Separation processes - crystallisation, flocculation/sedimentation**
- **3-d micrography of human eye**
- **Sampling from ice cores**
  - pollen, micro-meteorites, mammoths
- **Objects trapped in amber**

Other particle analysis:

- **Combustion processes & liquid atomisation**
  - water droplets, aerosols, clouds, snow and fog
- **Insect swarms**

HOLOMAR

Conclusions

The Grid is not a “magic bullet” – still need to think about how best to use computational and storage resources.

Can Grid be integrated into the data flow of the whole experiment, rather just the number-crunching bit?

Important note is not that the Grid makes it easy to “scavenge” idle resources, but rather that it makes cooperation to avoid wasting them in the first place much easier.