VICTORIA UNIVERSITY OF WELLINGTON

Department of Computer Science

PO Box 600
Wellington
New Zealand

Tel: +64 4 471 5328
Fax: +64 4 495 5232
Internet: Tech.Reports@comp.vuw.ac.nz

The Design and Implementation of MetVUW Workbench Version 1.0

Aaron Roydhouse   Linton Miller
Eric K. Jones   James McGregor

Technical Report CS-TR-93/7
November 1993

Abstract

MetVUW Workbench is a multimedia database for retrieval and display of historical meteorological data, including laser disc video imagery, digital satellite imagery, text, hydrological data, and numeric fields from the European Centre for Medium-Range Weather Forecasting. Existing components of MetVUW Workbench allow this information to be retrieved by date and time, by full text search, or by scanning the laser disc.

In developing MetVUW Workbench, interesting issues arose in five areas: interface design, memory organisation and data storage, management of the laser disc player, audio presentation of data, and use of publicly available software packages. This document reports on our experiences in the design and implementation of version 1.0 of MetVUW Workbench.

Author Information

Aaron Roydhouse and Linton Miller are Masters students within the Department of Computer Science at Victoria University of Wellington. They implemented MetVUW Workbench version 1.0 during the summer of 1992–93. Aaron Roydhouse was responsible for the overlay computation, full text search, online help, audio output, and false colouring components. Linton Miller was responsible for the user interface, data storage, indexing, data compression, and the device driver for the laser disc player.

Eric Jones is a lecturer with the Department of Computer Science at Victoria University of Wellington and a member of the department's artificial intelligence group. James McGregor is a senior lecturer in meteorology with the Institute of Geophysics at Victoria University of Wellington.

1 Introduction

MetVUW Workbench is a multimedia database of historical meteorological data. The system is designed to assist weather forecasters and other meteorologists, and to serve as a test bed for research into intelligent retrieval mechanisms. Version 1.0 of the system was completed in April 1993.

MetVUW Workbench allows retrieval and display of the following kinds of data: laser disc video imagery and corresponding digital satellite imagery, time-tagged text descriptions, hydrological data, and numeric fields from the European Centre for Medium-Range Weather Forecasting (ECMWF). Examples of numeric fields include pressure, temperature, relative humidity, wind speed, and relative vorticity. The data currently covers a fixed region including New Zealand, the southwest Pacific, the Tasman Sea, and Eastern Australia. Data can be retrieved by date and time, by full text search, or by scanning the laser disc. A number of mechanisms are provided for manipulating and displaying data once they have been retrieved, including false-colouring of satellite imagery, calculating contour overlays, scrolling text windows, and animating laser disc video. An experimental audio output device is also provided.

The current system accomplishes the following design objectives:

1. *Seamless access to multiple sources of data*

   It is possible to use any of the several search mechanisms provided to retrieve all data available at a given time. For example, a user can locate an image of interest using the laser disc player and then create contour displays of numerical fields at the time of that image. Alternatively, a user can locate a document of interest using full text search and then have instant access to the numerical data and laser disc imagery for the time period to which the document refers. A user can also retrieve images or numerical data by date and time and then start a laser disc animation at that point.

2. *Ease of Use*

   MetVUW Workbench is intended to provide straightforward access to the operations most commonly required by meteorologists. We therefore provide a graphical point and click interface that makes it very easy for users to carry out searches and to specify options for displaying data.

   This approach contrasts radically with the design philosophy of other existing meteorological databases such as the McIdas system employed at the Australian Bureau of Meteorology. McIdas provides a powerful command line interface with much greater functionality than MetVUW Workbench, but forecasters apparently find it very difficult to use in practice.

3. *Speed of Retrieval*

   We intend later versions of MetVUW Workbench to be actively used by operational forecasters. Fast retrieval is vital to these users, as they operate under severe time pressure. The current version of MetVUW Workbench runs on a SparcStation 10 model 30 and takes at most two seconds to execute any command from a user to retrieve or display data.

Version 1.0 of MetVUW Workbench serves several groups of users. First, research students in the Institute of Geophysics at Victoria University of Wellington use MetVUW Workbench as a tool for meteorological research. Second, within the Department of Computer Science, the system is used as a test bed for research and development of intelligent mechanisms for retrieving historical meteorological data. Two of us are actively engaged in the design of a mechanism for retrieving past weather situations from descriptions of high-level features of the meteorological processes and systems (Jones and Roydhouse, 1993).

The remainder of this report is structured as follows. Section 2 gives an overview of the design issues that arose in constructing version 1.0 of MetVUW Workbench. Section 3 describes the underlying architecture of the system. Section 4 presents the user interface design. Section 5 discusses organisation and storage of the data. Section 6 describes interaction the laser disc player. Section 7 discusses audio presentation of data. Section 8 describes the contribution of publicly available software packages. Section 9 concludes with directions for future work.

2 Overview

A number of interesting problems and issues arose in the design and implementation of MetVUW Workbench. These divide into five broad categories:

1. *User Interface*

   MetVUW Workbench is operated by an easy-to-use graphical user interface that provides seamless access to multiple data types (satellite imagery, text, numerical fields, and so forth), multiple search mechanisms (date and time, full text search, laser disc), and multiple devices (computer screen, laser disc player, audio). Managing this complexity gives rise to several interesting issues of interface design, two of which are as follows:

   (a) Coordination of multiple data types

   After retrieving data of one type, a user can immediately access all other types of data occurring at the same time. For example, having located a satellite image by browsing the laser disc, pressure and wind data for that date and time are immediately available. The user interface provides "synchronisation" buttons for this purpose that align different data types by their time of occurrence; see section 4.7.

   (b) Control of the laser disc player

   The laser disc player is operated by means of the same point and click interface used to drive the rest of the system, as opposed to employing an external device such as a remote control. MetVUW Workbench provides a control panel that relies heavily on the analogy to a video tape player, although it provides a greater range of functions; see section 4.3.

2. *Memory Organisation and Data Storage*

   To facilitate access to multiple types of data using a variety of search mechanisms, an appropriate memory organisation is required. In particular, given data of one type occurring at some point in time, it must be possible to retrieve all other types of data for the same time period. For example, given a satellite image, it must be possible to construct contour overlays for numerical fields such as mean sea level pressure. The indexing and retrieval strategies that MetVUW Workbench employs for this purpose are described in section 5.1. MetVUW Workbench also provides a number of data compression mechanisms, one for each data type, and allows compressed and uncompressed data to be freely intermixed; see section 5.2.

3. *Communication with the Laser Disc Player*

   No Unix-based device driver for the laser disc player was available, so we designed our own. Two problems arose in its design. The first problem was to design a high-level interface to the device driver that is more useful than the raw device commands. Our command set is described in section 6.1.

   Second, certain laser disc commands such as searches and repeats generate asynchronous return values, so the device driver has to be able to accept these return values at any time and process them in a sensible fashion. The complications this gives rise to, and our solution to them, are described in section 6.2.

4. *Audio presentation of data*

   In cooperation with the Electricity Corporation of New Zealand, we implemented an experimental audio output device that outputs audio tones of varying frequencies corresponding to times series of numeric data. Whilst browsing the laser disc, the user can simultaneously listen to an audio track representing time series of numeric data such as lake inflow rates. The audio driver is discussed in section 7.

5. *Use of publicly available software*

   A common fault of many software developments is to try and reinvent the wheel by writing all components of the project from scratch. Designing a new text search package, a database system, or a data compression algorithm would be interesting tasks in themselves, but they were not central to our goals. Instead, we used existing packages wherever possible to achieve speedy development and allow us to concentrate on the novel components and aims of MetVUW Workbench. The packages we used are freely available on the Internet. We discuss the contribution of these packages in section 8.

3 Architecture of MetVUW Workbench

This section describes the internal architecture of MetVUW Workbench. We first describe the organisation of the elements of the system, then provide brief details on the implementation.

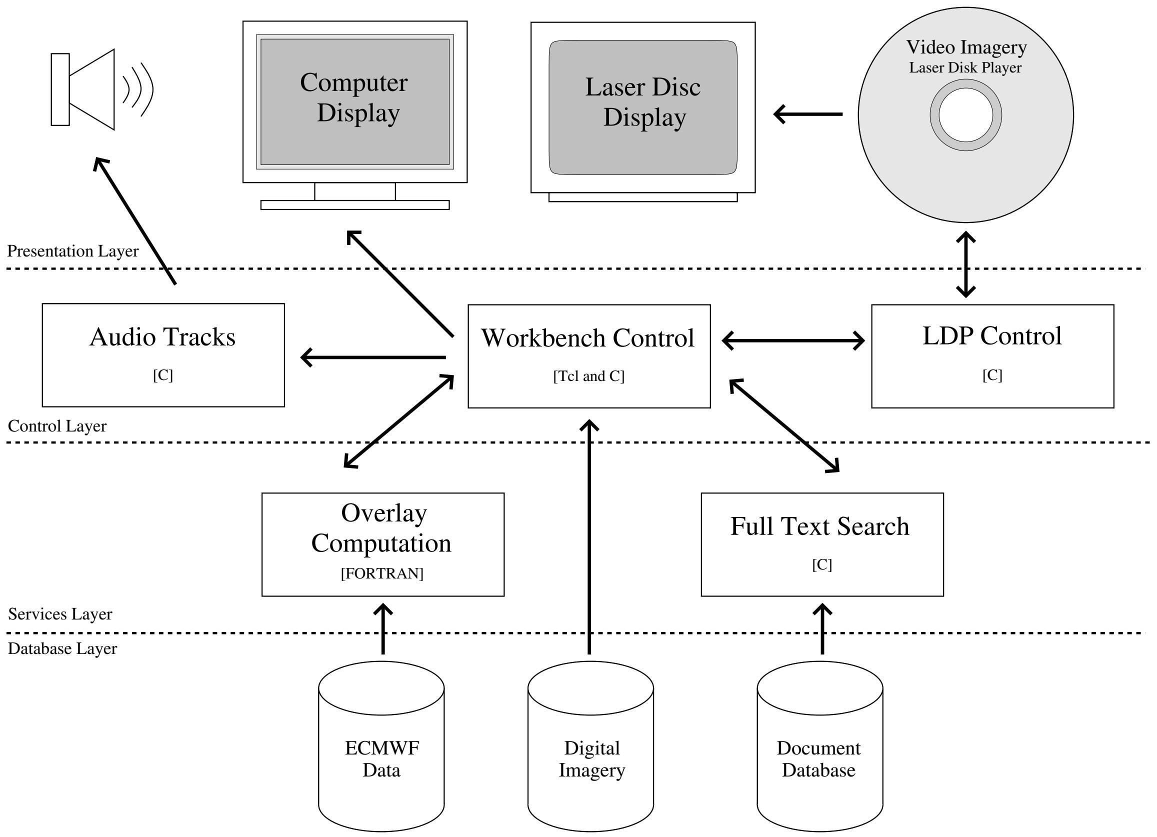

Figure 1: MetVUW Workbench Organisation

The MetVUW Workbench architecture consists of four layers, as shown in figure 1:

- The **Presentation Layer** is the user's view of the system. A high-resolution computer monitor displays digital imagery, overlays of contours and wind barbs, and a graphical user interface. Laser-disc video imagery is displayed on a second monitor. Audio is generated on the host computer's speaker.
- The **Control Layer** supervises the laser disc player and audio devices, and brings together requested digital satellite images, text documents, and overlays for presentation on the computer display.
- The **Services Layer** computes contours and wind barbs and carries out searches of the document database.
- The **Database Layer** manages large collections of numerical data, digital images, and text documents.

Because of the scope of tasks required for the MetVUW Workbench implementation, no one language is suitable. Version 1.0 of MetVUW Workbench is implemented in Tcl/Tk, C, and FORTRAN. The graphical user interface and core of MetVUW Workbench is implemented in Tcl/Tk. The drivers for the laser disc player and audio output are implemented in C. The overlay computation library is derived from code written by Dr. Crispin Marks and implemented in FORTRAN. This code includes a contouring algorithm which uses the St. Andrews Cells method and thus correctly contours saddle points (Sabin, 1986). Other packages for image compression (JPEG), full text search (LQ-text), database management (NDBM), and Tcl/Tk itself are implemented in C. The Tcl code acts as the glue that brings these components of the implementation together to form a complete application, as discussed in section 8, below.

The satellite images available to the system are stored in two different formats: on hard disk in a digital format, and on laser disc in an analogue video format for purposes of animation. The satellite images are available for each hour and show a region covering New Zealand, the southwest Pacific, the Tasman Sea, and eastern Australia. ECMWF data are available for every twelve hours and at 2.5 degree grid-points. Some of the data are available for 14 atmospheric levels from surface to the 10mb level, including pressure, temperature, potential energy, geopotential, vertical wind velocity, horizontal wind velocity, and relative humidity. Other ECMWF data are by their nature available at only one specific level, including two-metre dew-point, two-metre temperature, mean sea level pressure, surface pressure, surface geopotential, and ten-metre wind. A number of “derived” fields are pre-calculated from the ECMWF data, including wind speed, surface wind speed, 300mb thickness, 500mb thickness, potential vorticity, relative vorticity, residual force, and heating rate. The text document collection includes brief synopses of a subset of the satellite images, and articles from the newsletters of the Meteorology Society of New Zealand.

The development platform for MetVUW Workbench is a SparcStation 10 model 30 running SunOS 4.1.3. The laser disc player is a Sony 3600. The languages and compilers used were Tcl 6.7, gcc 2.3.3, Sun FORTRAN 77 1.4p6, and Texinfo 2.16. The library versions used were X11R5, Tk 3.2, JPEG 4.0, SunOS 4.1 NDBM, LQ-text 1.12gamma, and TkInfo 0.3.

4 Interface Design

The user interface of MetVUW Workbench integrates a number of different mechanisms for retrieving and displaying meteorological data. The system currently supports retrieval and display of laser disc video imagery, digital satellite imagery, hydrological data, ECMWF data, derived numeric parameters, and time-tagged text descriptions. These resources are all accessible using the graphical user interface shown in figure 2. We next discuss each component of the interface.

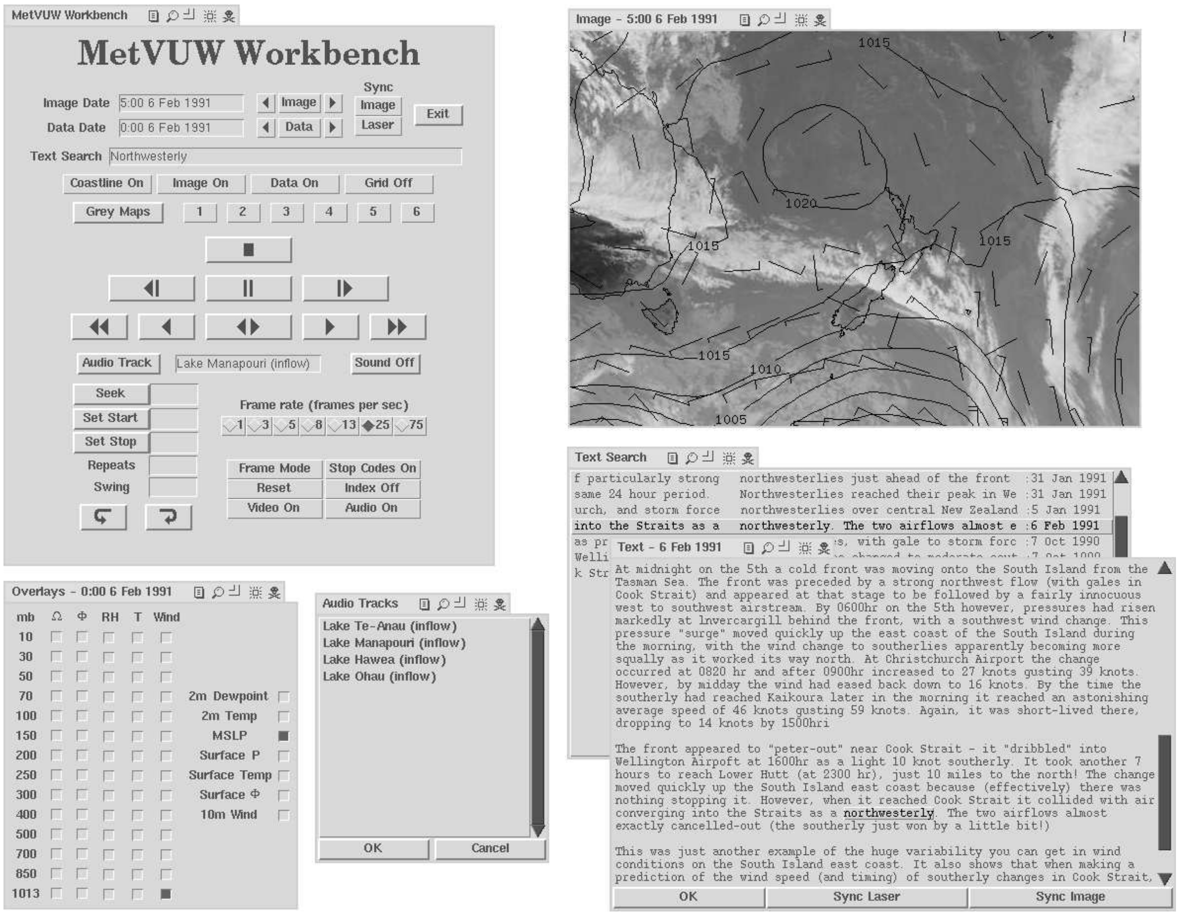

Figure 2: The MetVUW Workbench graphical user interface

4.1 Retrieval by Date

A date and time entered in to the "Image Date" entry box, shown in figure 3, will cause the corresponding satellite image to be retrieved. If the requested image is missing for some reason, then the image for the nearest available hour is retrieved. In this case the Image Date entry box is changed to reflect the actual date and time of the displayed image.

Retrieval by image date also selects the closest ECMWF data set, providing it is available within six hours of the image date. The "Data Date" entry box is changed to contain the date and time of the selected ECMWF data. In a similar fashion, ECMWF data can be retrieved by entering a date and time in the "Data Date" entry box; if available, an image within two hours of the data date will be retrieved.

Stepper buttons step forward and backward, image by image or data set by data set, synchronizing the other type of data at each step.

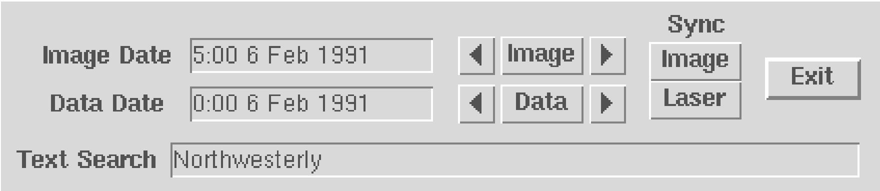

Figure 3: Retrieval by date or key phrase

4.2 Retrieval by Text Search

A key phrase entered in the "Text Search" entry box, shown in figure 3, causes a window to pop up that lists all occurrences of this phrase in the online document archive. Each line of the text search window gives the context of an occurrence of the key phrase and the date of the document in which it occurs. Selecting any of these occurrences causes the whole document to be presented in a separate window.

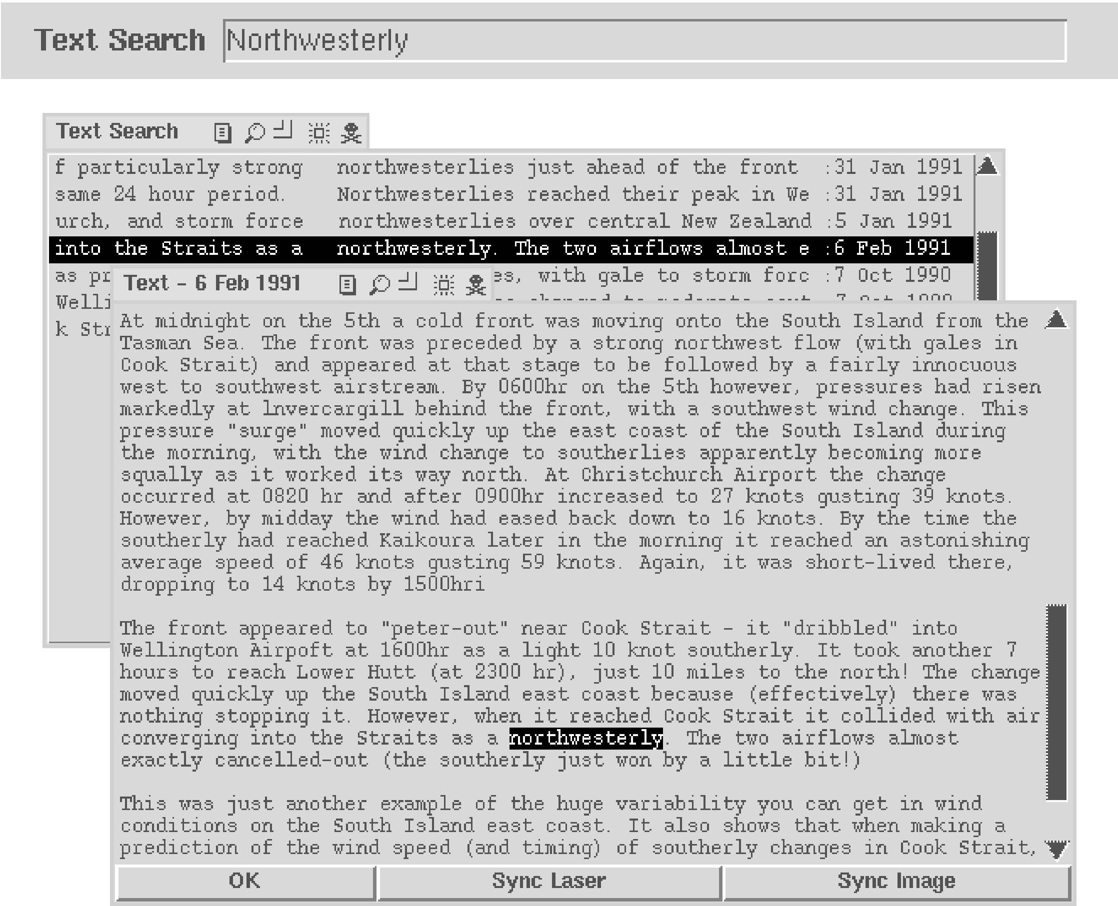

Figure 4: Retrieval by text search

Figure 4 shows the text search entry box containing the search phrase "Northwesterly." The results of the search are shown in the text search window, partially obscured by a window containing the full text of one of the documents.

Each window that contains a document also has buttons for retrieving laser disc or digital imagery occurring at that time. These buttons are described in section 4.7.

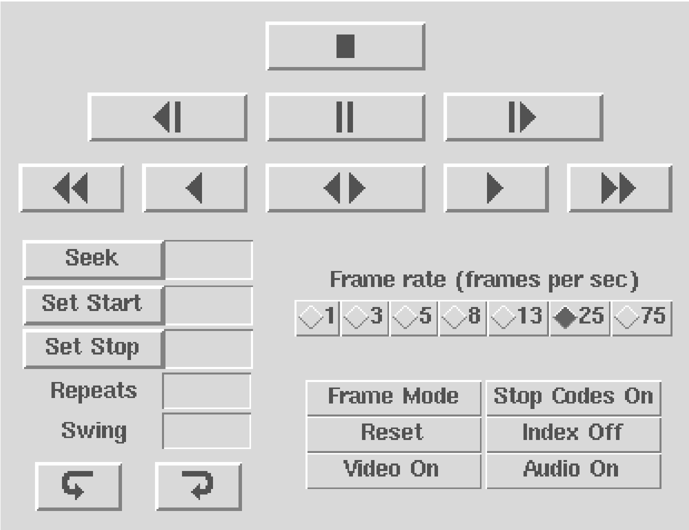

Figure 5: Laser disc player controls

4.3 Laser Disc Player Controls

MetVUW Workbench provides a standard set of laser disc controls for playing and searching through the laser disc, as shown in figure 5. The three rows of buttons at the top are as follows: stop, frame back, still frame, frame forward, rewind, play backward, play other direction, play forward, and fast forward. The right-hand middle row of buttons controls the frame rate for playback operations. At the bottom right is a set of toggle buttons for features of the laser disc player. These features are as follows: seek by frame or chapter, stop at stop codes, reset the player, display frame index on screen, display video, and play audio. On the middle left, the "Seek" button displays the frame whose number is in the adjacent entry box. On the bottom left are entry boxes and buttons for controlling repeat and swing loops. Repeat loops play a sequence of satellite images a number of times in row. Swing loops are like repeat loops except that on each cycle, the sequence is played forwards then backwards, and a larger number of frames can be played in one direction than the other on each cycle. The "Set Start" and "Set Stop" buttons determine the frames to be repeated. The "Repeats" box sets the number of times to play a repeat loop. The "Swing" entry box determines the number of frames by which swing loop iterations are offset. The pair of buttons at the very bottom left initiate repeat and swing loops; the left-hand button plays loops backwards, and the right-hand button plays loops forwards.

4.4 The Image Window

The Image Window displays digital satellite imagery and overlays of contours and wind barbs. The region covered includes New Zealand, the southwest Pacific, the Tasman Sea, and Eastern Australia. Data for many atmospheric levels are available; see section 4.6. For example, figure 6 shows an infra-red satellite image overlayed by contours for the mean sea level pressure and "wind barbs" indicating the speed and direction of the surface winds.

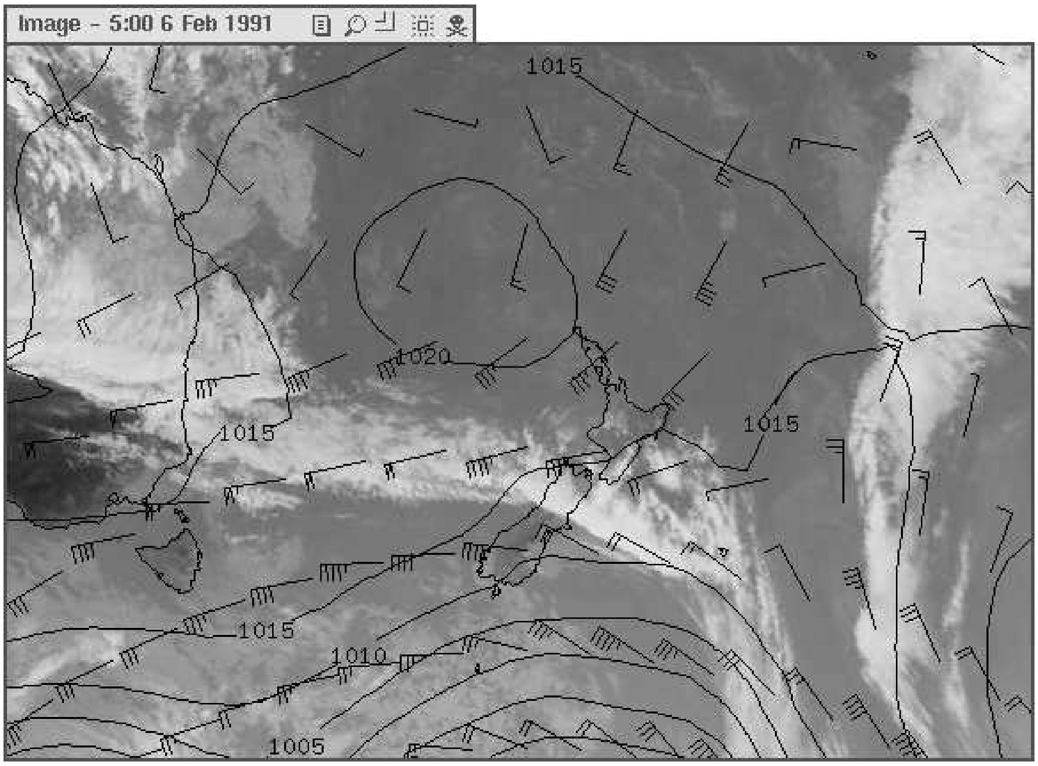

Figure 6: Image window showing mean sea level pressure and wind barbs

4.5 Image Controls

Figure 7 shows the interface for the image controls. Several toggles and options affecting the image window are available:

- **Coastline On/Off** toggles an overlay showing an outline of land masses.
- **Image On/Off** specifies whether or not digital satellite imagery is displayed.
- **Data On/Off** specifies whether or not data overlays are displayed.
- **Grid On/Off** toggles an overlay showing lines of latitude and longitude.

Figure 7: Image controls

The lower row of buttons control the false-colouring facilities. Twelve options are available: six colour and six grey-scale. Each option is a piecewise linear function that maps segments of data values to ranges of colours. A toggle button is provided to switch between colour and grey-scale option sets; buttons labelled “1” to “6” select an option.

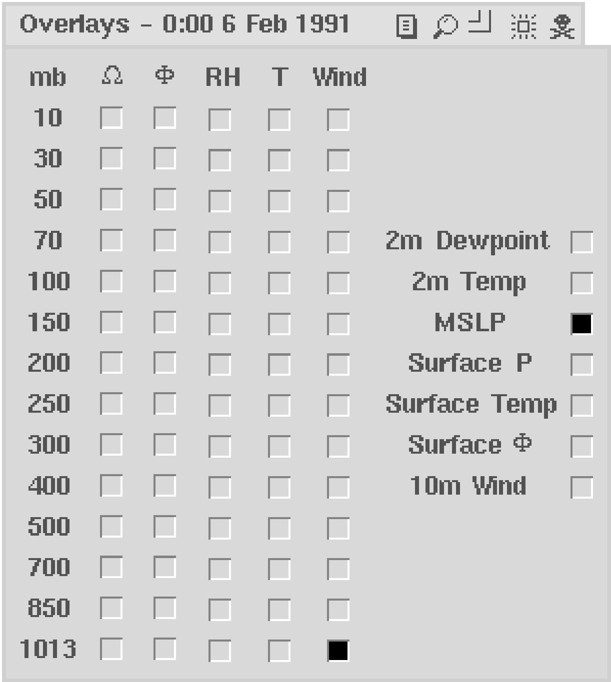

Figure 8: ECMWF overlays

4.6 ECMWF Overlays

A separate window, shown in figure 8, allows the user to specify contour and wind-barb overlays for the current data date. The rightmost column contains data fields that exist at only one atmospheric level. On the left, is an array of data fields. Each column of the array refers to a different data field; each row refers to a different level of the atmosphere. Following standard meteorological practice, pressure is used as a measure of height.

In figure 8, mean sea level pressure (MSLP) and wind speed and direction for 1013mb are currently selected. The overlays created for this selection are shown on the image in figure 6.

4.7 Synchronisation Buttons

Synchronisation buttons, shown in figure 9, align data for a given time with other types of data for the same time.

Image synchronisation buttons, shown at the top of figure 9, align digital image data with laser disc data and vice versa. Selecting sync "Laser" causes the laser disc player to display the video image corresponding to the currently selected digital image. These two images are identical except that one is stored in an analogue video format and the other in a digital format. Similarly, selecting sync "Image" cause the image window to display the digital image corresponding to the video image currently displayed by the laser disc player.

Figure 9 also shows a text window containing a document retrieved as the result of a full text search. Synchronisation buttons at the base of the text window trigger retrieval of the laser disc and digital images occurring at the time of the document.

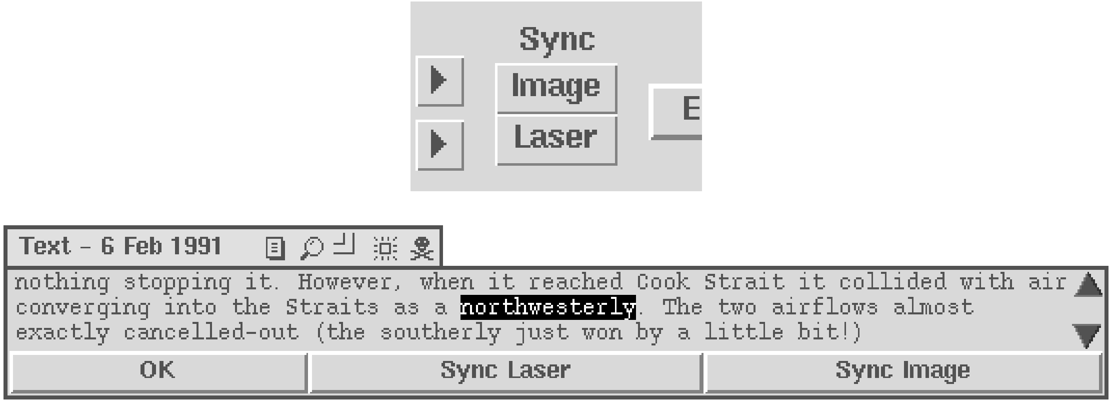

Figure 9: Synchronisation buttons

4.8 Audio Tracks

Audio tracks of hydrological data can be played synchronously with the laser disc imagery as described in section 7, below. A selection box allows the user to select the data series they wish to hear, whilst that generation of audio can be toggled on and off at any time; these controls are shown in figure 10.

Figure 10: Audio controls and track selection

4.9 Help Facility

MetVUW Workbench provides a hypertext help facility, shown in figure 11. Help information is organised as a hierarchy of menus, with hypertext-style cross references. Clicking on any window or button of the computer display will bring up a new window displaying a hypertext node that contains appropriate help information. This information may include cross references and a menu of related topics. Clicking on a menu item or cross reference will cause that topic to be displayed.

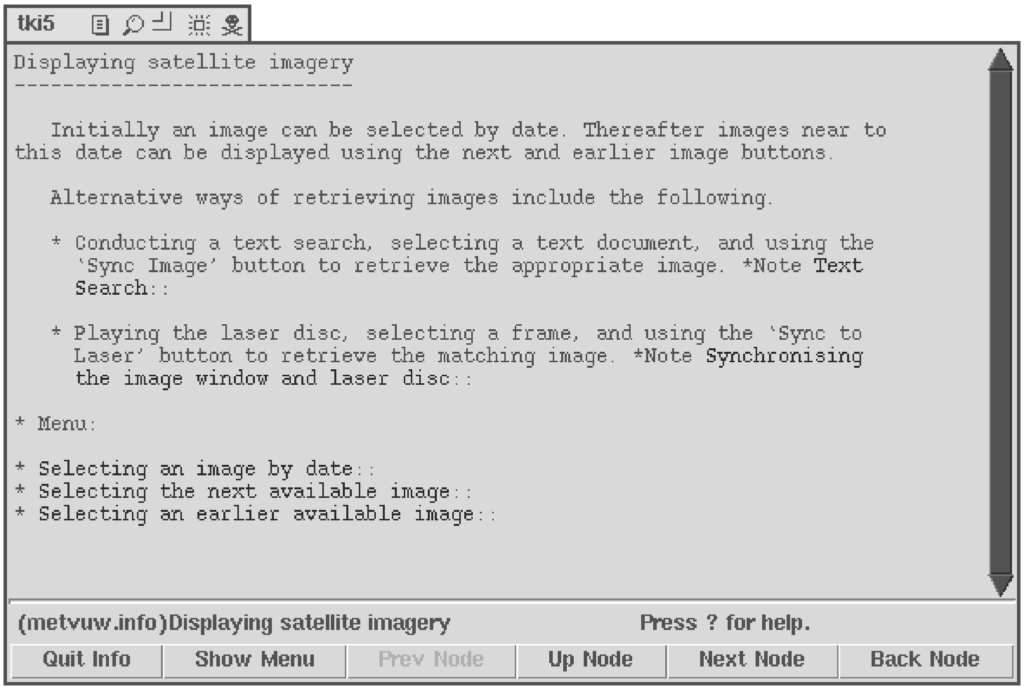

Figure 11: Help facility

5 Memory Organisation and Data Storage

5.1 Data Indexing and Synchronisation

To facilitate access to multiple types of data using a variety of search mechanisms, all data are indexed by the date and time of their occurrence. Most of these indices are bidirectional: not only can data be located by date and time, but given a datum, the date and time of the datum can be recovered. For example, a table is provided that records the date and time of each frame number of the laser disc player. Using this table, the current frame number displayed by the laser disc player can serve as in index to retrieve the date and time of the image. A similar table can then be used in the other direction to retrieve the corresponding digital image for that date and time.

In addition to these indices, every non-textual datum provides access to its immediate predecessor and successor in the historical record of data of that type. These "relative" indices are used for stepping forwards and backwards through the record. Further indices to textual data are also compiled for full text search.

In more detail, the following indices are provided:

1. *Laser disc frame index*

   A table indexed by frame number contains a date and time for each frame. This table is used for synchronising other data sources with the image displayed by the laser disc player.

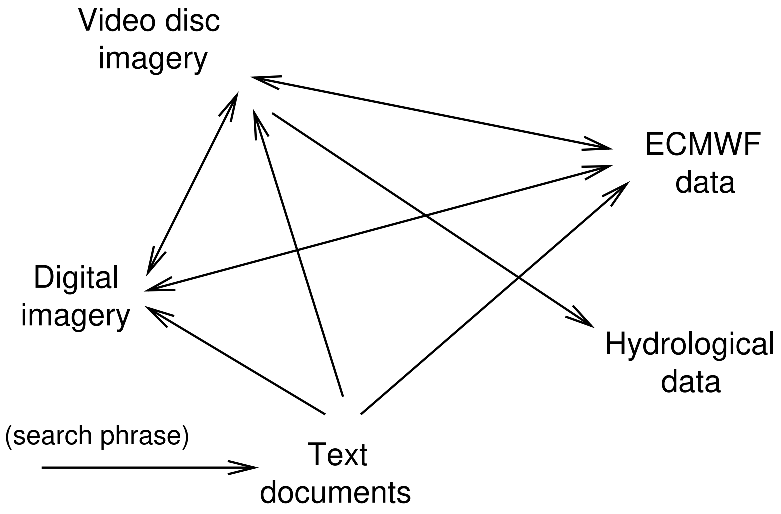

Figure 12: Data connectivity provided by version 1.0 of MetVUW Workbench

2. *Date and time index*

   A table indexed by date and time permits access to a number of data sources. For each date and time, the table includes the following items

   (a) The *frame number* for the image on the laser disc that depicts the closest date and time available.
   (b) The *frame numbers* for the closest available laser disc images for later and earlier times.
   (c) An *image filepath* for the file that contains the digital image.
   (d) *Image filepaths* for the closest available images for later and earlier times.
   (e) A *data filepath* for the set of files containing numeric data.
   (f) *Data filepaths* for the closest available data sets for earlier and later times.

3. *Filename index*

   A table indexed by filename contains the date to which each file refers. This index is used to recover the date and time from a filename. For example, if the date and time index is used to obtain a filepath for the closest earlier image for a given date and time, then the date and time of that earlier image can be identified using the filename index.

4. *Array of hydrological data*

   Hydrological data are coupled to laser disc imagery by storing it in a file that can be directly indexed by frame number (see section 7).

5. *Text concordance*

   A concordance records all occurrences of each word used in the collection of text documents and is constructed automatically by the full text search system. The concordance is used to locate text phrases in the document collection.

The freedom to move from one source of data to another that MetVUW Workbench provides is shown in figure 12. Interconnectivity between the different data sources is nearly complete, with one important exception: the existing implementation does not permit access to text by date and time, so text cannot be retrieved by way of retrieving other kinds of data. This needs to be fixed.

5.2 Data Compression

To reduce storage requirements, meteorological data can be stored in compressed form. The appropriate form of compression varies according to data type. Presently we employ two compression algorithms, one for images, and another for numerical data.

Images are compressed using an algorithm called JPEG, which is specifically designed for lossy compression of still images. Images compressed with JPEG appear almost indistinguishable from the original even when compressed to a tenth of their original storage requirements.

Numeric data are compressed using Lempel-Ziv coding, a general-purpose lossless compression algorithm. The effectiveness of Lempel-Ziv depends of the type of data being compressed. For our numeric meteorological data we achieved 10–20% compression. In the future we hope employ domain specific compression techniques for numeric data to improve on this figure.

6 Communication with the Laser Disc Player

6.1 High Level Interface

The laser disc driver provides a high-level, easily callable interface to all the features of the Sony 3600 laser disc player. The driver is naturally device-specific, but should convert with only minor modification to other models of Sony player.

The laser disc player is a device capable of displaying video frames and sound from a laser disc on a television monitor. Images can be displayed still frame, or played back at a variety of speeds, either forwards or backwards. A sequence of frames can be looped repeatedly. Frames are sequentially numbered, and the player can randomly access any frame from its number. The player can play both PAL and NTSC format discs.

Although the laser disc player can play back at a variety of speeds, each speed requires sending a different command to the player. Rather than providing a separate set of playback commands for each possible speed, the driver maintains a current *frame rate* for the player. The frame rate determines the speed of all playback commands. It is the job of the driver to translate a high-level playback call, in conjunction with the current frame rate, into the appropriate laser disc player command. Actual frame rates are determined by the format of the disc (PAL or NTSC).

The repeat loops provided by the player are only capable of looping a fixed cycle of frames forwards or backwards at normal playback speed. The driver enhances these loops to provide repeat loops or swing loops played at the current frame rate. Swing loops are like repeat loops except that on each cycle, the sequence is played forwards then backwards, and a larger number of frames can be played in one direction than the other on each cycle.

Other capabilities of the driver include support for control of multiple players, full error handling, and the ability to connect to the laser disc player via TCP/IP or direct serial connection.

The complete command set of the driver is listed in figure 13.

6.2 Implementation

All the high level commands listed in figure 13 are provided as a library of function calls that are linked with the main program. These function calls generate commands that are sent to the laser disc player. Communication with the laser disc player is via a 9600 baud serial line. All commands are single byte codes, possibly followed by multiple byte arguments. Normally, the player returns

Connection functions
Open
Close

Status functions
Get disc id
Get frame address
Get chapter address
Get frame rate
Get status
Is playing

Playback functions
Set frame rate
Stop
Still
Frame advance
Start
Scan
Set memory position
Memory search
Search
Repeat
Swing

Control functions
Reset
Clear
Eject
Video on
Video off
Audio on
Audio off
Set frame mode
Set chapter mode
Stop codes on
Stop codes off
Index on
Index off
Audio channel 1
Audio channel 2
Motor on
Motor off
Eject enabled
Eject disabled
Noise reduction on
Noise reduction off

Figure 13: Function calls provided by the laser disc player driver

either an acknowledgement byte signalling that the command was received and is acceptable, or an error byte. On status queries, however, instead of returning an acknowledgement, the player returns a stream of bytes encoding the player status. The player also asynchronously sends bytes when the laser disc is ejected manually from the player or the end of a repeat loop is reached. These asynchronous bytes may arrive at any time except in the middle of a stream of bytes returned from a status query.

Whenever asynchronous bytes arrive, the operating system raises an interrupt to notify the driver. The driver then catches the interrupt, reads the asynchronous bytes and takes the appropriate action. It is possible that only one interrupt will be raised for a group of asynchronous bytes, so the driver must be able to deal with many outstanding bytes any time an interrupt is raised.

The arrival of asynchronous bytes can cause problems. Normal command processing happens as shown in figure 14, where a command is sent, and the player response then read. However, it is possible that an asynchronous byte will be sent by the player at the same time as the driver is about to send a command to the player; see figure 15. Thus the first byte on the serial line after the command is not the response from the command but the asynchronously received byte. The driver copes with this problem using the method shown in figure 16.

A limitation of the driver is its inability to notice some changes in the state of the laser disc player. Currently there is no feedback from the driver to any caller when a repeat loop is completed, nor when the disc is ejected. There is also no feedback from the laser disc player to the driver when the player reaches the end of a disc and stops playing. The only way to detect that the player has reached the end of the disc is to constantly poll the status of the laser disc player—not a desirable option because of the very high overhead this incurs.

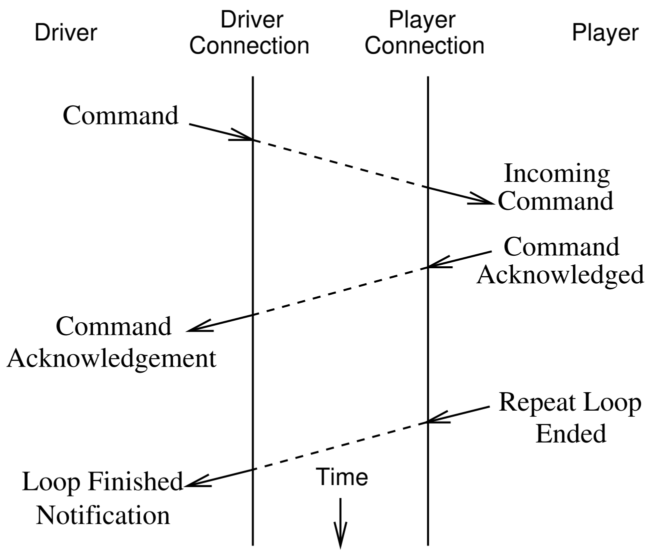

Figure 14: Normal communication with the player

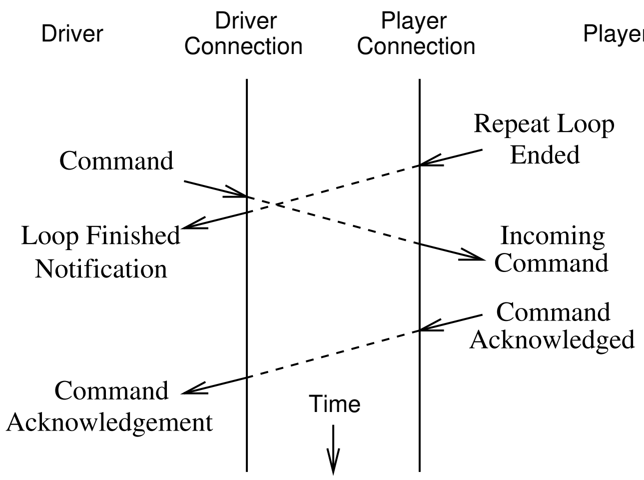

Figure 15: Asynchronous return overlapping a player command

```
When an interrupt is received:
  For each outstanding asynchronous byte
    If it indicates the completion of an iteration of a loop,
       Decrease the number of iterations left, and if necessary,
       do another iteration.
    Else if it indicates that the disc has been ejected (manually)
       Reset player state to reflect the lack of a disc

To send a command:
  Disable interrupts
  Send command
  Repeat
    Read 1 byte
    If it is an asynchronous byte, queue it to process later
  Until byte read is not asynchronous
  Read any more bytes expected as the result of the command
  Process queued asynchronous bytes
  Re-enable interrupts
```

Figure 16: Handling asynchronous bytes

7 Audio Presentation of Data

The audio driver plays numeric data "in sync" with the laser disc player imagery. The Electricity Corporation of New Zealand provided the MetVUW project with data series for lake inflow rates. While the laser disc player operates, the hydrological data are retrieved and mapped to an audio tone. Different data values give rise to an audio tone at a different frequency. As each laser disc frame triggers matching hydrological data the tone is ramped from the previous level to the new level, as shown in figure 17.

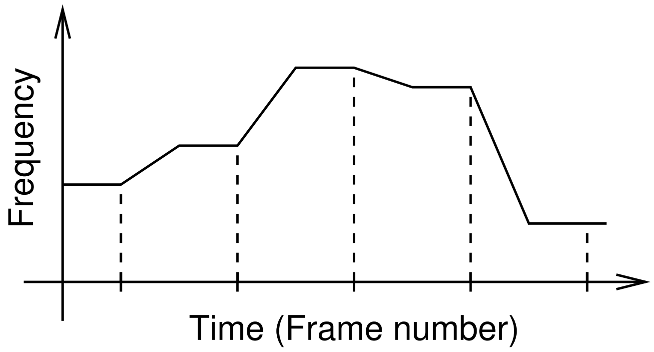

Figure 17: Frequency versus time for the mapping of hydrological data to audio frequencies, showing frequency ramping between data.

8 Use of Publicly Available Software

Version 1.0 of MetVUW Workbench was constructed very quickly, and is a substantial application; nevertheless, it is already stable and relatively bug-free. Instrumental in this success was the use of innovative tools contributed to the Internet. We now discuss the role of these packages play MetVUW Workbench.

1. *Graphical User Interface*

   Tool Command Language (Tcl) is an interpreted and extensible programming language (Ousterhout, 1991). The most popular set of extensions to Tcl is a graphics tool kit, called Tk. Tk is a library for implementing of graphical applications for X-windows. We used the Tcl/Tk combination to implement the graphical user interface for MetVUW Workbench. Because Tcl/Tk code controls both the user interface and the top-level program control, we are able to closely couple the state of the program and the interface. For example, the date and time of the current image displayed by the graphical user interface is directly linked to the program state variable storing the current date and time of the image. Whenever this variable is changed by the system, the graphical user interface is automatically updated.

2. *Full Text Search*

   MetVUW Workbench employs a package for full text search called LQ-text. LQ-text is able to quickly locate all instances that match a key phrase in a large corpus of indexed free-format text. LQ-text supports simple "stemming," which means, for example, a search phrase containing "northwesterly" will also match "northwesterlies" in the text. MetVUW Workbench uses LQ-text to retrieve documents that match a key word or phrase from a collection of meteorological documents. After such a search the user can immediately retrieve other sources of data for the period to which a document refers.

3. *Image Compression*

   One of the challenges of the MetVUW Workbench project is dealing with high volumes of satellite imagery. Hourly satellite images for one day requires 6Mb of storage in raw form, while a year of imagery requires over 2Gb of storage. Ordinary lossless compression algorithms such as Lempel-Ziv only reduce this requirement by about 30%. We therefore employ the lossy image compression algorithm specified by the Joint Photographic Experts Group or JPEG (Wallace, 1991; Pennebaker and Mitchell, 1993). Images compressed with JPEG appear almost indistinguishable from the original, even when compressed to a tenth of their storage requirements. Using JPEG, a year of imagery requires only 200Mb of storage.

4. *Help Facility*

   MetVUW Workbench Version 1.0 is already a large application that requires a manual and an effective online help facility. We wrote the manual using GNU'S Texinfo system (Stallman and Chassell, 1993). The Texinfo program automatically encodes the text of the manual in two ways: as a word-processed document, and in a hypertext format suitable for use by the GNU Info hypertext system. The hypertext version of the manual is built into the graphical user interface of MetVUW Workbench using a Tcl/Tk interface package call TkInfo designed for this purpose. Appropriate sections of the manual are indexed in terms of the corresponding windows and buttons of the user interface. Clicking on any window or button of the user interface pops up a window containing the appropriate section of the manual. Users can then follow further cross references to retrieve related information. MetVUW Workbench uses the TkInfo user interface to the GNU Info hypertext system to display the online help.

9 Conclusions and Future Work

Our initial efforts in developing MetVUW Workbench have proved highly successful. Version 1.0 of MetVUW Workbench is already a powerful and easy-to-use package for retrieval and display of historical meteorological data. The system is in active use within the Institute of Geophysics at Victoria University of Wellington, and has attracted considerable interest from the Meteorological Service of New Zealand.

We plan to extend the present system in two ways. First, we intend to add an intelligent retrieval system that allows forecasters and other meteorologists to retrieve past weather situations by specifying high level descriptions of them. The system is to serve as a "memory amplifier" that allows users to rapidly locate historical situations of interest. In particular, forecasters should be able to use the system to quickly retrieve situations that are similar to the current one in meteorologically significant respects, providing them with an additional source of information to supplement the output of sometimes equivocal numerical models

Second, we intend to extend and improve MetVUW Workbench's existing facilities for retrieval and display of data in a number of ways, including the following:

1. *Improve the database management system.*

   Presently, indices to data are pre-built and static, which make it difficult to add or remove data. We plan to construct facilities that make it easy to add or delete data. To solve this problem, we intend to switch to an extensible RDBMS called Postgres as a substrate to future versions of MetVUW Workbench (Stonebraker and Rowe, 1986). This change will also facilitate the development of intelligent retrieval techniques.

2. *Provide customisable false colouring and contouring.*

   Currently, false colouring schemes are presets defined in configuration files, contours can only be constructed at preconfigured intervals, and contouring is limited to the horizontal plane. We plan to allow users to specify their own false colouring schemes using the graphical user interface. Users should also be able to specify different intervals for contouring, and request contouring of vertical or oblique planes through the three dimensions of meteorological data.

3. *Derive numerical fields online.*

   At present, all data fields must be pre-computed, including fields such as wind velocity that we derive ourselves from the component fields. These derived fields exacerbate the high data storage requirements of MetVUW Workbench. We plan to allow some of these fields to be calculated on demand, allowing storage requirements to be significantly reduced.

4. *Develop techniques for domain dependent compression of numeric data.*

   We already use JPEG, a domain dependent compression algorithm, to compress image data. We plan to also introduce domain dependent techniques for compressing numeric data.

5. *Improve the laser disc interface.*

   Most user interface controls for the laser disc player adhere closely to the analogy to a video player remote, but some advanced features do not fit this model. In particular, repeat loops and swing loops of frames are currently specified by frame number; it should also be possible to specify them by date and time. Also, the addition of a time-line view of the laser disc would enhance the user's model of the data, making it possible to select animation periods and particular dates by point-and-click.

6. *Make use of frame grabbing technology.*

   Satellite images on laser discs are stored in an analog format, so they are not readily computer-readable. However, frame-grabber technology allows a video signal to be reinterpreted as a digital image. With the addition of frame-grabber hardware to the SparcStation, we could animate satellite imagery, directly on the computer display. This would eliminate the need for a separate video monitor and for storage of satellite imagery in a digital format if corresponding video imagery is available on laser disc. It will also be possible to false-colour these animations and to calculate contour overlays in real time. (Note that the slow data rate of the hard disk drive precludes rapid animation of digital imagery.)

7. *Improve organisation of document collections.*

   The current system allows full text search of document collections, but provides only a primitive mechanism for organising and displaying the documents that are retrieved. We expect the size of the document collection to increase considerably in the near future, so this mechanism will soon be inadequate. We therefore propose to construct a more sophisticated facility for summarising and displaying the results of full text searches that is better suited to larger and more diverse document collections.

Acknowledgements

This research was supported in part by IGC research grants VUW/V212/451/RGNT/594/410A and VUW/V212/451/RGNT/594/448, and by a grant from the Electricity Corporation of New Zealand. Additional funds were supplied by the Department of Computer Science and the Institute of Geophysics at Victoria University of Wellington. Thanks to Dr. Crispin Marks for provision of contouring code. Kim Rutherford implemented a prototype for the audio presentation of hydrological data. Finally, we would like to thank staff at the Meteorological Service of New Zealand for helpful advice at various stages in the design of the system.

References

Jones, E. K. and Roydhouse, A. (1993). Intelligent retrieval of historical meteorological data. In *Proceedings of a Workshop on Artificial Intelligence and the Natural World*, Melbourne, Australia. Forthcoming.

McGregor, J. and Jones, E. K. (1993). MetVUW: A multimedia teaching aid. In *The First International Conference on Computer-Aided Learning and Distance Learning in Meteorology, Hydrology, and Oceanography*, Boulder, Colorado.

Ousterhout, J. K. (1991). An X11 toolkit based on the Tcl language. In *Proceedings of the Winter USENIX Conference*, pages 105–115. USENIX Association.

Ousterhout, J. K. (1994). *Tcl and the Tk Toolkit*. Addison-Wesley. Forthcoming.

Pennebaker, W. B. and Mitchell, J. L. (1993). *JPEG Still Image Data Compression Standard*. Van Nostrand Reinhold.

Sabin, M. (1986). A survey of contouring methods. *Computer Graphics Forum*, 5:325–340.

Stallman, R. M. and Chassell, R. J. (1993). *The Texinfo Manual, Version 2.18*. Online documentation.

Stonebraker, M. and Rowe, L. (1986). The design of Postgres. In *Proceedings of the 1986 SIGMOD Conference on Management of Data, Washington, D.C.* ACM Press.

Wallace, G. K. (1991). The JPEG still picture compression standard. *Communications of the ACM*, 34(4):30–44.

Appendix 1: MetVUW Video Disc

MetVUW

What is MetVUW video disc?

MetVUW video disc is a high quality interactive video disc and is a superb resource for research and training in meteorology. It was first demonstrated at an International Workshop on Cyclonic Weather Systems in the Tasman Sea and South West Pacific at Wellington, New Zealand in January 1993, and holds all of the laser video imagery used by MetVUW Workbench version 1.0.

The disc contains over 16 000 separate hourly images from the infrared channel on board the Japanese Geostationary Meteorological Satellite (GMS). The images are for an area covering the Eastern seaboard of Australia, the Tasman Sea, New Zealand and the South West Pacific. The imagery covers a period from 1 July 1990 to 21 July 1992. The disc also contains some ECMWF analysis fields and composites of sea surface temperature.

When played at normal video playback speeds, the video disc gives an excellent overview of over two years of meteorology in eleven and a half minutes. The sequence gives a superb insight into the complexities of the mid-latitude weather systems that affect the region.

Moreover, when used interactively and at slower speeds, MetVUW video disc enables a detailed examination of a vast number of individual weather systems on the synoptic and meso scale.

Who produced MetVUW video disc and why?

MetVUW video disc was produced by the Institute of Geophysics at Victoria University of Wellington, New Zealand for research into the dynamics of weather systems in the Tasman Sea region. The disc is also used extensively in graduate and undergraduate teaching at Victoria University.

How can I play MetVUW video disc?

MetVUW video disc can be played on any PAL or dual PAL/NTSC video disc player. The video output from the player can be plugged into any television, monitor, or television projector that confirms to PAL standards.