www.czasopisma.pan.pl PAN www.journals.pan.pl

ARCHIVES
of
FOUNDRY ENGINEERING

10.24425/afe.2020.133341

Published quarterly as the organ of the Foundry Commission of the Polish Academy of Sciences

ISSN (2299-2944)
Volume 2020
Issue 4/2020

5 – 12

1/4

The Use of the CUDA Architecture to Increase the Computing Effectiveness of the Simulation Module of a Ceramic Mould Quality Forecasting System

M. Hojny [a], **K. Żaba** [b], **T. Dębiński** [a, *], **J. Porada** [a]
[a] AGH - University of Science and Technology, Department of Applied Computer Science and Modeling,
ul. Czarnowiejska 66, 30-054 Krakow, Poland
[b] AGH - University of Science and Technology, Department of Plastic Processing and Metallurgy of Non-Ferrous Metals
al. A. Mickiewicza 30, 30-059 Kraków, Poland
* Corresponding author. E-mail address: tdebinsk@metal.agh.edu.pl

Received 09.017.2020; accepted in revised form 21.04.2020

Abstract

This paper presents practical capabilities of a system for ceramic mould quality forecasting implemented in an industrial plant (foundry). The main assumption of the developed solution is the possibility of eliminating a faulty mould from a production line just before the casting operation. It allows relative savings to be achieved, and faulty moulds, and thus faulty castings occurrence in the production cycle to be minimized. The numerical computing module (the DEFFEM 3D package), based on the smoothed particle hydrodynamics (SPH) is one of key solutions of the system implemented. Due to very long computing times, the developed numerical module cannot be effectively used to carry out multi-variant simulations of mould filling and solidification of castings. To utilize the benefits from application of the CUDA architecture to improve the computing effectiveness, the most time consuming procedure of looking for neighbours was parallelized (cell-linked list method). The study is complemented by examples of results of performance tests and their analysis.

Keywords: Computer simulation, casting, CUDA, Smoothed particle hydrodynamics

1. Introduction

Precise casting in lost-wax ceramic moulds enables geometrically complicated shapes of aircraft parts, such as turbine blades, vane segments, turbine nozzle rings, and casings to be manufactured. These components are made of nickel superalloys, and due to their operating conditions (high pressure, operation temperature above 1000°C) they are called critical parts of jet engines. The production process of aircraft engine parts is very complex, and their ultimate quality is influenced by many material, technological, and organisational factors. Multi-layer ceramic casting moulds are made in a multi-operational process of applying subsequent ceramic layers onto wax pattern sets, de-waxing, washing, drying and baking the mould. Layers are applied either manually or by robots. After baking, moulds are warmed, preheated to an adequately high temperature and filled with liquid metal in vacuum conditions. The cost of material, usually very expensive nickel and cobalt superalloys, and the subsequent post-casting treatment operations account for a substantial part of manufacturing costs (even up to 70%) of critical aircraft engine parts. Each of the mentioned individual processes may induce defects, which can be either repairable or not. Therefore, the accurate quality inspection of semi-products is very important, so that wax pattern sets and multi-layer ceramic casting moulds not meeting the requirements – which can lead to manufacturing faulty castings – are eliminated already at an early stage [1,2]. The occurrence of surface and voluminal defects is a problem limiting the qualitative development of production of aircraft engine parts with the precise casting process. Periodic occurrence of defects with unidentified nature leads to disqualification of from 20 to 40% of castings from a single heat, despite using advanced manufacturing technologies and product quality inspection methods. This issue is extremely important, because it concerns large series industrial production of precise castings of aircraft engine parts such as blades, vane segments, turbine nozzle rings and casings, which have a decisive effect on the flight safety, and therefore cannot be faulty. Despite many years of research and experiments it has not been possible to define all reasons of defect occurrence, even though the manufacturing processes have been specifically monitored. This fact is characteristic of the difficulties and determines the direction of research and its scope. Taking into account the complexity of factors influencing the quality of individual unit processes, semi-products, and consequently a finished product and ineffective, currently used, methods of quality inspection, it seems reasonable to develop a solution based on knowledge in the form of a modular, dedicated system for quality forecasting and aiding optimisation of the precise casting process, to ensure that competitiveness of aircraft part production is achieved and maintained. The hybrid methodology of modelling casting solidification processes, potentially taking into account the industrial geometry of the mould featuring a variable thickness on its cross-section is one of key components of methods and procedures to aid engineering, developed as part of the proposed solution. A key component of the developed methodological solution is a spatial numerical model based on the smoothed particle hydrodynamics (SPH) [3,4,5]. The intensive development of computer technologies, accompanied by the continuous growth of the computing capacity, enables more and more complex physical processes to be effectively modelled. As processing big data is needed, it has become necessary to use parallel computing. Thanks to the CUDA technology created by NVIDIA, it has become possible to use graphic processors for general purpose computing. Applying this technology has allowed data processing to significantly accelerate and it has found application in many areas such as data enciphering, where the authors of paper [6] obtained about 87 times faster computing for algorithm AES-128 in the parallel version performed on a GPU than on a CPU. For similar applications the authors of paper [7] obtained acceleration of about 20x. Another area of application of fast GPU computing is medical imaging, where – as shown in papers [8,9] – the computing was accelerated about 1000x depending on an image processing algorithm. Another area requiring processing big data, in particular for computing performed in the real time, is the simulation of flows, e.g. in medicine [10] or mechanics [11], where acceleration of computing was achieved at a level of about 20x, depending on the applied algorithm. In neural networks GPU computing is successfully used [12,13]. In this area the achieved acceleration of computing compared to CPU is from about 18x to 48x. There are solutions allowing the CUDA technology to be applied in conjunction with the finite element method, where as presented in paper [14], acceleration of computing is from 12x to 50x, depending on the algorithm executed. On the basis of the foregoing literature, one can state that the use of the CUDA architecture will allow the computing capacities of the numerical computing module, being part of a comprehensive system for ceramic mould quality forecasting, to be effectively increased.

2. Modular Quality Forecasting System

The Modular Quality Forecasting System (Fig. 1) comprises the Measurement Data Acquisition Subsystem, consisting of the 3D Scanning and Thermovision Modules, being key modules of the system, the Tomography Module applied for creating quality standards at new production start-ups, and the Vision System Module and the Profilometry Module, being complementary modules and the Computing and Control Subsystem with the Data Processing and Analysis Module and the Numerical Simulation Module. The acquisition modules include modern equipment and accessories for non-destructive testing, using state-of-the-art innovative solutions, focused on high accuracy, repeatability and stability of the measuring process, at their small variability, associated with advanced devices and IT software.

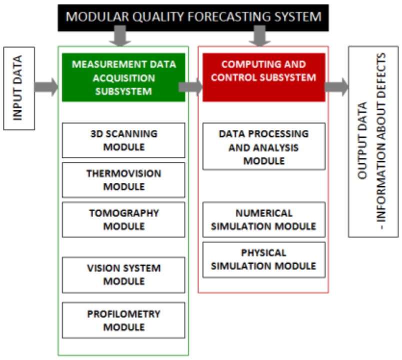

Fig. 1. Schematic diagram of the Modular Quality Forecasting System

2.1. Vision System Module

High resolution cameras fitted with specialist filters are the primary components of the module for vision testing of the wax pattern quality. The transparency effect is used for the quality inspection of wax patterns. The analysed pattern is illuminated, the image is recorded, and in the next step it is interpreted using dedicated software. It results in matching the contour of the recorded image of the wax pattern to a CAD file of the standard and determining material defects present in the pattern, which cannot be effectively inspected visually.

2.2. Thermovision Module

The main component of the thermovision module is a FLIR T640 thermographic camera with a measurement range from -40 to 2000°C, detector resolution 640x480, and accuracy ±2°C. The module's task is to analyse heating or cooling of ceramic moulds in order to determine the temperature field distribution, its difference from the standard shows faulty areas of the mould. In addition, after correlating the temperature field distribution measurement results with measurement results of wall thickness, obtained with 3D scanning or industrial computer tomography, it is possible to assess this parameter in terms of quality and quantity in the real time in a process line. Fig. 2 shows an example of a thermogram of a multilayer ceramic mould during cooling. The diverse temperature in various areas of the mould indicates its various thickness and potential defects such as porosity.

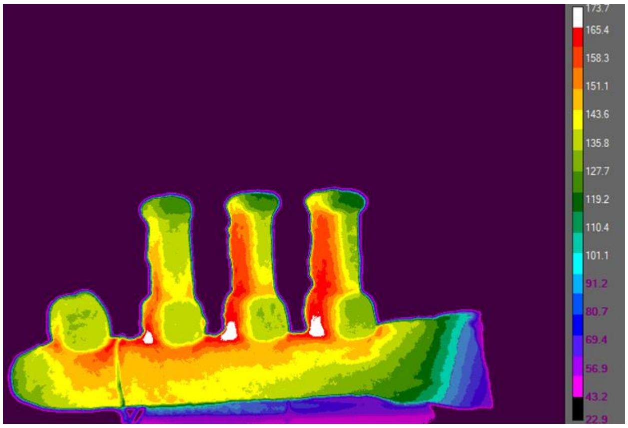

Fig. 2. Thermogram of a multilayer ceramic mould during cooling

2.3. Scanning Module 3D

The 3D scanning module is equipped with an ATOS 3D scanner by GOM with the GOM Inspect software. The measurement accuracy is 0.02 mm. The main purpose of the module is to scan wax patterns and wax pattern sets in order to determine their shape and dimensions, and to measure the mould thickness by overlaying mould and pattern set scans. Fig. 3 presents the view of a wax system for jet engine blades along with the mould gating system, after scanning to transform it to a digital form. The digital data are recorder in the *.STL file format, where areas are described with a huge number of objects, that is triangles.

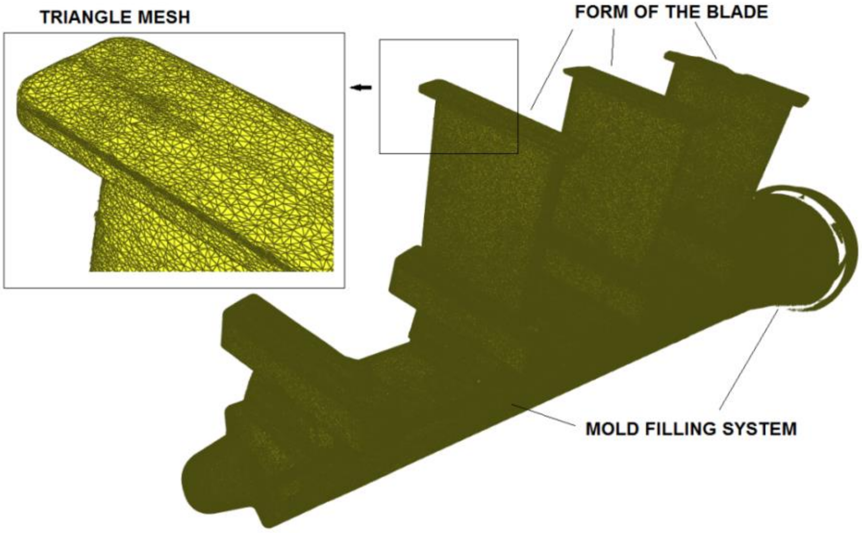

Fig. 3. A wax system after the scanning to digital form

After digitalisation of the ceramic form and calibrating the obtained area relative to the scan of the wax system, a real mapping of the mould thickness used in industry may be obtained (Fig. 4). The analysis of the obtained map of the thickness leads to the observation that it features a variable thickness on its cross-section, reaching the minimal value of 3.519 mm in the area of the filling system. The maximum value of 14.378 mm was determined in the area of the jet engine blade mould.

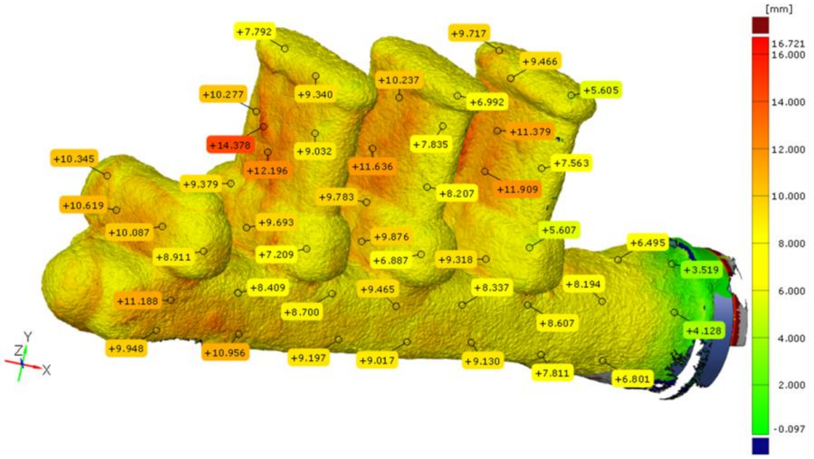

Fig. 4. Map of real thickness of the mould obtained in mapping the 3D scans of the ceramic mould and the wax system

The consideration for the real geometrics of the mould in numerical calculations will allow the process of filling and solidifying in terms of locally variable changeable heat conditions to be simulated, which as a result will allow potentially defect-prone areas in the cast volume to be more precisely estimated. On the other hand, if the finite element method (FEM) is used to simulate the solidification process, an approach like this allows any inconveniences related to the generation of spatial 3D meshes (digitizing the actual mould geometry) to be avoided. For the SPH method, the actual mould geometry described with points is automatically generated on the base of scanning results for the external surface of the mould and the standard wax set.

2.4. Tomography Module

The main component of the tomography module is a high-resolution phoenix v|tome|x L 450 microtomograph, for both 3D and 2D x-ray tests. The task of the tomography module is to test the mould layer quality, discontinuities inside the casting and within parts inaccessible for a 3D scanner, to determine exact percentage mould porosity, mould wall thickness and to create quality standards. Fig. 5 presents an example of mould image from the tomograph (longitudinal section), along with the selected porosity marked with blue in the picture.

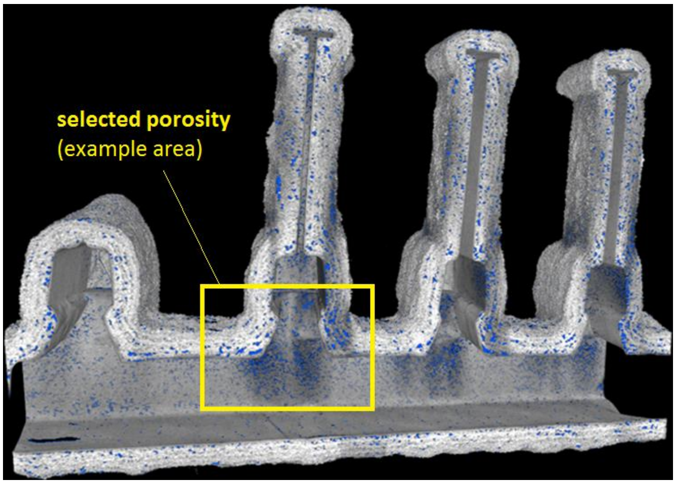

Fig. 5. Ceramic mould section obtained from the tomograph, along with the selected porosity (blue areas)

2.5. Profilometry Module

The main component of the profilometry module is a Wyko NT930 optical profilometer. The module allows the wax pattern and casting surface quality to be precisely analysed, and the first mould layer next to the pattern to be analysed for new product launches.

2.6. Data Processing and Analysis Module

The Data Processing and Analysis Module is one of the three modules comprising the Computing and Control Subsystem. This module enables measurement data processing to be automatized with the dedicated software, and standard and advanced process data analysis to be performed. The primary purpose of the module is to create correlation between the final properties of ceramic moulds and their thickness and temperature during mould making processes, which involve an elevated temperature (e.g. after taking out of an autoclave).

2.7. Physical Simulation Module

A Gleeble 3800 thermo-mechanical simulator is playing here a key role, as this is a fundamental tool for simulating the selected manufacturing processes. The essence of physical simulation is the reconstruction – in the laboratory conditions – of changes in temperature, strain and stress to which the material is subjected in an actual industrial process. Instead of full-size materials processed, small samples made of the same material that is used in the production process are used in the tests. The evaluation of mechanical properties of samples, which are subjected to various physical simulation variants, is the basis for developing guidelines to enable the optimal parameters of the process line equipment operation to be determined [15].

2.8. Numerical Simulation Module

The numerical simulation module uses the original DEFFEM 3D software to model and simulate the precise casting process [15]. It uses an innovative method of simulation with actual data referring to the mould wall thickness as a result of the 3D scanning module operation. Fig. 6 shows an example of simulation result in the form of temperature field distributions on the selected longitudinal and transverse sections of a jet engine blade.

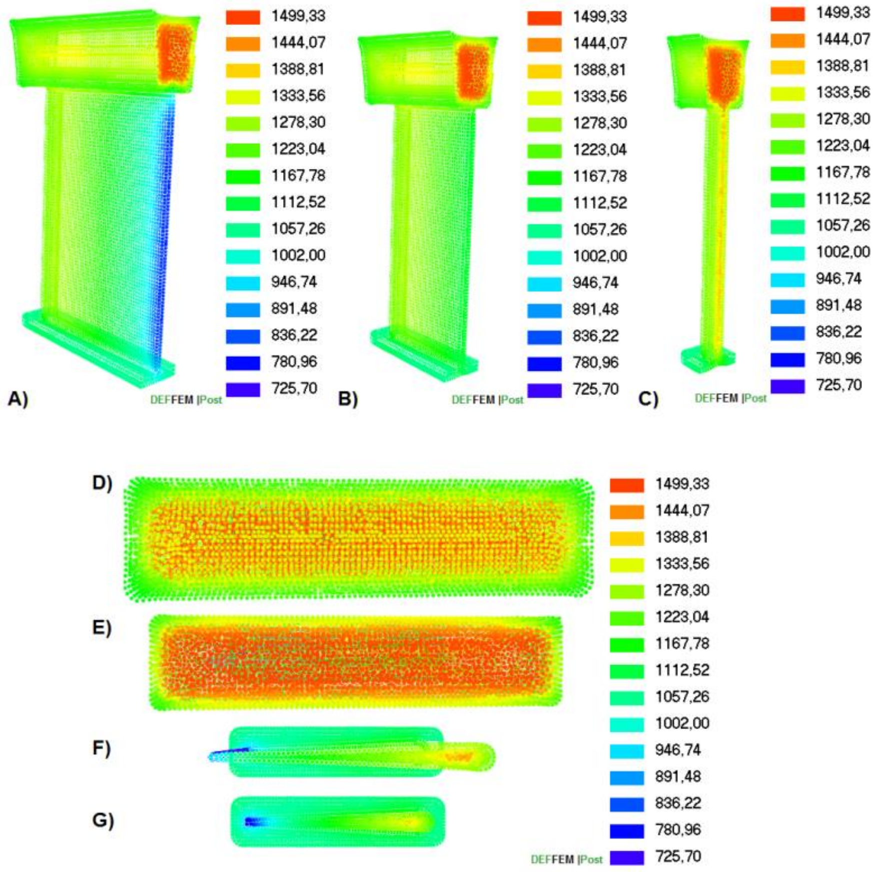

Fig. 6. Example of temperature field distribution on the selected sections of a jet engine blade (A-C section with plane X, D-G section with plane X)

3. Application of the CUDA architecture to increase the computing efficiency – examples of test results

The numerical implementation of the smoothed particle hydrodynamics computing module is executed as part of the original DEFFEM 3D package, developed for over ten years [15]. In accordance with the idea of the SONATA 2 project, financed by the National Science Centre in 2012-2018, the developed software is the foundation of the scientific technique focused on aiding high-temperature processing of steels and alloys [15].

Implementation work of the mathematical model was performed in the NetBeans IDE 8.0 program environment using the Fortran language. Fig. 7 presents the computing diagram of the SPH solver, being part of the DEFFEM 3D package. One of the components significantly influencing the computing effectiveness is the neighbouring particle searching algorithm. The current version of the SPH solver is based on the Cell-linked list (CLL) algorithm. This method operation algorithm divides a three-dimensional solution domain into cubic cells with set dimensions, where each particle is assigned to a specific cell. The essence of this CLL method efficiency is in looking for neighbouring particle in neighbouring cells only.

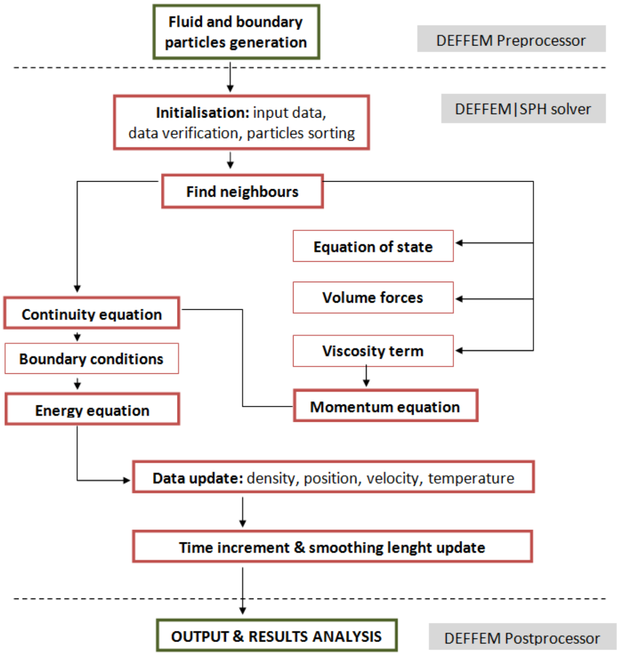

Fig. 7. The computing diagram of the SPH numerical computing module of the DEFFEM 3D package

Executing the CLL algorithm sequentially on the CPU processor, when at the same time the particle number is large, significantly affects the computing efficiency. As a result, the developed numerical tool significantly restricts possibilities for practical application in industrial conditions. To increase the computing effectiveness it was proposed to use the universally available and relatively cheap computing technology – the CUDA. This technology allows many computing cores of a GPU graphic processor to be used for general purpose computing. Computing for a parallel version of the CLL algorithm was executed using a Tesla T4 graphic card by NVIDIA. Its most important parameters are presented in Table 1.

Tests were conducted for two variants of the called CUDA kernel:

- Variant 1: where the grid consists of 16 1D blocks, each block has 32 threads (marking grid (16,1,1) and block (32,1,1)),

-Variant 2: where the grid consists of 16 1D blocks, each block has 1024 threads (marking grid (16,1,1) and block (1024,1,1)).

The neighbour search algorithm test was based on a steel solidification simulation in a mould with dimensions 1x1x1m (Fig.8). The test material was C45 grade steel. The liquidus T_l and solidus T_s temperatures of the tested steel are 1412°C and 1494°C, respectively. The thermal properties necessary for numerical calculations was predicted by using commercial JMatPro software [15].

Table 1.
Parameters of the Tesla T4 card used for computing

Global memory	15080 MBytes
Max. number of threads per block	1024
Max. number of threads per multiprocessor	1024
CUDA cores	2560
Multiprocessors	40

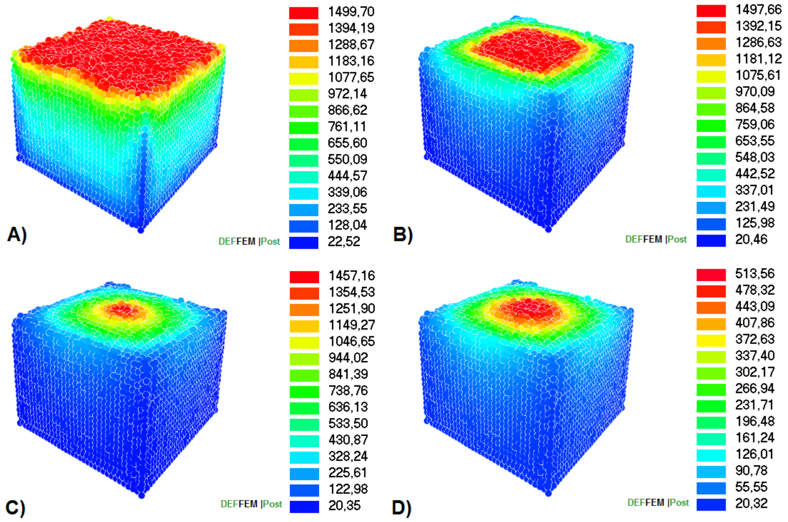

Fig. 8. The temperature distribution for selected stages of ingot solidification process simulation (test simulation)

Summary of the obtained computing times for a single integration time step and for the example of particle number is presented in Table 2. Findings in the graphic form are presented in Fig. 9. On this graph, the computing time for the GPU is indicated on the vertical axis on the left hand side, while for the CPU on the vertical axis on the right hand side.

Table 2.
Summary of the computing times of the CLL algorithm for the CPU and GPU, [s]

PARTICLE NUMBER	CPU	GPU - 1st variant	GPU - 2nd variant
100000	40.8	1.8	0.6
200000	180.6	7.8	1.8
300000	444.6	17.4	4.8
400000	832.8	30.6	9
500000	1450.8	48	14.4
600000	2397	69	21.6
700000	3525.6	93.6	30
800000	4789.8	122.4	40.2
900000	6216	154.8	52.8
1000000	7908	190.8	66

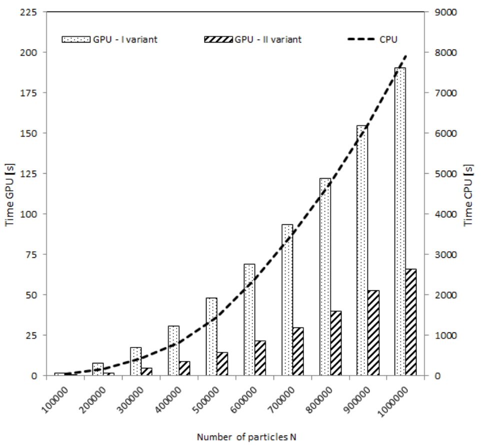

Fig. 9. Summary of the computing times of the CLL algorithm for the CPU and GPU

Analysing the obtained results (Table 2, Fig. 9) one can observe that parallelization of the CLL algorithm responsible for looking for the nearest particles allowed this code fragment to be computed on average about hundred times faster. One should also note that this difference increases as the solution domain size increases (number of particles).

4. Conclusions

This article presents a modular system of ceramic mould quality forecasting, which enables a faulty ceramic mould to be eliminated from a process line before the casting operation. The numerical computing module is one of intensively developed areas of the system. Due to the applied high precision of the smoothed particle hydrodynamics (SPH) method the computing times significantly limit the suitability of this software in the industrial conditions. To minimise the mentioned limitation a modern computing architecture CUDA was applied to increase the effectiveness of the neighbour search algorithm (CLL method). The conducted performance tests showed that the use of the CUDA architecture enables the computing time to be shortened about hundred times for a single computing step compared to the execution time with the main CPU processor. As a result, the suitability of the developed numerical tool has been significantly improved, and thereby numerical simulations of ceramic mould casting or casting solidification in moulds have been performed more effectively.

Acknowledgements

1. The tests and the development of the foundations of the DEFFEM 3D package, being the foundation of the scientific technique focused on high-temperature steel processing financed by the National Science Centre (2012-2018), Decision number: DEC 2011/03/D/ST8/04041
2. Developmental research and industrial tests of the SPH solver financed by the National Centre for Research and Development (2016-2018), Applied Research Programme PBS3/A5/54/2016

References

[1] Atlas of Wax Pattern Defects. Investment Casting Institute, 2004.

[2] Atlas of Shell Defects. Montvale, Investment Casting Institute, 2004.

[3] Monaghan, J.J. (1992). Smoothed particle hydrodynamics. *Annual Review of Astronomy and Astrophysics*. 30, 543-574.

[4] Monaghan, J.J. (2005). Smoothed particle hydrodynamics. *Reports on Progress in Physics*. 68, 1703-1759.

[5] Cleary, P.W., Ha, J., Prakash, M. & Nguyen, T. (2006). 3D SPH flow predictions and validation for high pressure die casting of automotive component. *Applied Mathematical Modelling*. 30(11), 1406-1427.

[6] Abdelrahman, A., Fouad, M. (2017). High performance cuda aes implementation: a quantitative performance analysis approach, Computing Conference, July, 18-20 London, UK.

[7] Manavski, S.A. (2007). Cuda compatible gpu as an efficient hardware accelerator for aes cryptography, IEEE International Conference on Signal Processing and Communications (ICSPC 2007), 24-27 November, Dubai, United Arab Emirates.

[8] Kalaiselvi, T., Sriramakrishnan, P. & Somasundaram, K. Survey of using GPU CUDA programming model in medical image analysis. *Informatics in Medicine*. 9, 133-144.

[9] Shi, L., Liu, W., Zhang, H., Xie, Y. & Wang, D. (2012). A survey of GPU-based medical image computing techniques. *Quant Imaging Med Surg*. 2(3), 188-206.

[10] Lee, H., Kim, B., Lee, K., Jung, H. (2016). Acceleration of computational fluid dynamics analysis by using multiple GPUs, Int'l Conf. Bioinformatics and Computational Biology. BIOCOMP'16, CSREA Press.

[11] Brandvik, T., Pullan, G. (2008). Acceleration of a 3D Euler solver using commodity graphics hardware,46th AIAA Aerospace Sciences Meeting.

[12] Huqqania, A.A.., Schikutaa, E., Yea, S., Chena, P. (2013). Multicore and GPU parallelization of neural networks for face recognition, Elsevier, International Conference on Computational Science, ICCS 2013.

[13] Uetz, R., Behnke, S. (2009). *Large-scale object recognition with CUDA-accelerated hierarchical neural networks*, Proceedings of the 1st IEEE International Conference on Intelligent Computing and Intelligent Systems (ICIS 2009).

[14] Akbariyeh, A. (2012). Large scale finite element analysis using gpu parallel computing, PhD Thesis.

[15] Hojny, M. (2018). *Modeling of Steel Deformation in the Semi-Solid State*, Switzerland.